University of Alberta

Characterization of the carboxyethylarginine synthase genes involved in clavulanic acid and 5*S* clavam metabolite biosynthesis in *Streptomyces clavuligerus*

by

Kapil Tahlan ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Microbiology and Biotechnology

Department of Biological Sciences

Edmonton, Alberta

Spring, 2005

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

0-494-08302-6

Your file Votre référence
ISBN
Our file Notre reference
ISBN

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Carboxyethylarginine synthase (CeaS) catalyzes the first reaction in the partially shared biosynthetic pathway leading to clavulanic acid and the 5*S* clavams in *Streptomyces clavuligerus*. Using Southern analysis, a second gene (*cea*S1) encoding CeaS was isolated, which is 73% identical at the nucleotide level to the original *cea*S2 gene from the clavulanic acid gene cluster. Along with *c7p* and *skn*, which encode a putative transcriptional regulator and sensor kinase, respectively, *cea*S1 is located next to other paralogous genes, which are also involved in the early stages of clavulanic acid and 5*S* clavam metabolite biosynthesis. This group of genes comprising *cea*S1 was designated as the paralogue gene cluster to distinguish it from the clavulanic acid and clavam gene clusters, which are also involved in the biosynthesis of the respective metabolites. Chromosome walking and Southern analysis of macro-restriction fragments of genomic DNA separated by pulsed field gel electrophoresis indicated that the above three genes clusters are not physically linked.

Single mutants defective in *cea*S1 and *cea*S2 were prepared, and were found to be affected to varying degrees in their ability to produce clavulanic acid and the 5*S* clavams, whereas the *cea*S1/*cea*S2 double mutant was completely blocked in the biosynthesis of the respective metabolites. In addition, preparation and analysis of mutant strains defective in *c7p* and *skn* revealed that they did not produce any detectable levels of the 5*S* clavams, but were unaffected in clavulanic acid production.

The nutritional regulation of *cea*S1 and *cea*S2 expression was analyzed and it was found that *cea*S1 was transcribed in complex soy medium only, whereas *cea*S2 was transcribed in both soy and defined starch asparagine media. As well, *cea*S1 transcription

was not affected in a mutant defective in the regulatory protein CcaR, whereas that of *cea*S2 was greatly reduced in the same mutant when compared to the wild type strain.

Overall, results suggest that different mechanisms are involved in regulating the expression of *cea*S1 and *cea*S2, and despite the structural and biosynthetic relatedness of clavulanic acid and the 5*S* clavams, the genes involved in their production reside in three unlinked gene clusters.

Acknowledgements

Much gratitude is owed to Dr. Susan E. Jensen for her patience and supervision, making the writing of this thesis possible. I would also like to thank Dr. Brenda K. Leskiw and Dr. Laura S. Frost, members of my supervisory committee, for their helpful advice and suggestions. I would like to thank Annie Wong, Cecilia Anders and Nathan Zelyas for their help, Jingru Li for being a good lab companion, Dr. Perrin H. Beatty and Dr. Hyeon Ung Park for their advice, and all other past members of the Jensen lab, as they had some part in the writing of this thesis.

I would like to express my gratitude to members of the Leskiw lab for their help with RNA analysis, and Rakesh Bhatnagar and Jack Scott from the Microscopy Unit, Department of Biological Sciences, University of Alberta, for contributing their expertise on confocal microscopy. In addition, I would like to thank members of the Digital Teaching Resources Laboratory and the Molecular Biology Service Unit at the Department of Biological Sciences, University of Alberta, for their assistance. It would also be appropriate to acknowledge the Department of Biological Sciences, University of Alberta, and the Alberta Heritage Foundation for Medical Research for funding.

I would also like to express my sincere gratitude to Dr. Dawn R.D. Bignell and all of my other friends for their companionship, and to my family who made all of this possible.

Table of Contents

Page

CHAPTER 1: INTRODUCTION.. 1

1.1 The *Streptomyces*.. 2
1.2 The β-lactam metabolites.. 5
1.3 Cephamycin C biosynthesis in *S. clavuligerus*.. 10
1.4 Biosynthesis of clavulanic acid and the 5*S* clavams.................................. 19
1.5 Regulation of β-lactam antibiotic biosynthesis in *S. clavuligerus*................. 38
1.6 Main objectives of this thesis.. 47

CHAPTER 2: MATERIALS AND METHODS.. 48

2.1 Materials.. 49
2.2 Bacterial strains and plasmids.. 50
    2.2.1 *Escherichia coli* and *Streptomyces* strains.............................. 50
    2.2.2 Plasmid vectors.. 50
2.3 Centrifugation and Harvesting of Cultures.. 50
2.4 *E. coli* procedures.. 62
    2.4.1 Culture conditions and media.. 62
    2.4.2 Preparation of *E. coli* glycerol stocks.................................. 62
    2.4.3 Preparation of electrocompetent *E. coli* cells.......................... 62
    2.4.4 Electroporation of DNA into electrocompetent *E. coli*................. 63
    2.4.5 Isolation of plasmid and cosmid DNA from *E. coli*..................... 64
    2.4.6 Preparation of colony blots.. 64
2.5 *Streptomyces* procedures.. 64
    2.5.1 Culture conditions and media.. 64
    2.5.2 Storage of strains.. 65
    2.5.3 Preparation of *S. clavuligerus* protoplasts.............................. 65
    2.5.4 Transformation of *S. clavuligerus* protoplasts.......................... 66

2.5.5 Introduction of DNA into *S. clavuligerus* by inter-generic conjugation........ 67
2.5.6 Plasmid DNA isolation........ 68
2.5.7 Isolation of gene replacement mutants........ 69
2.5.8 Fermentation of *S. clavuligerus*........ 70
2.5.9 Isolation of chromosomal DNA........ 71
2.5.10 Preparation of agarose plugs containing chromosomal DNA........ 72
2.5.11 Isolation of RNA........ 72
2.5.12 Preparation of cell-free extracts........ 74
2.6 Determination of levels of antibiotic production and growth........ 75
2.6.1 Growth determination........ 75
2.6.2 Bioassay for the detection of penicillin and cephamycin........ 76
2.6.3 Bioassay for the detection of clavulanic acid........ 76
2.6.4 Bioassay for the detection of alanylclavam........ 76
2.6.5 HPLC analysis for clavulanic acid and 5*S* clavam production........ 77
2.6.6 LC-MS analysis of clavulanic acid and 5*S* clavam production........ 78
2.7 DNA Analysis........ 78
2.7.1 Digestion, generation of blunt ended fragments and ligation reactions........ 78
2.7.2 *En bloc* digestion of *S. clavuligerus* DNA for pulsed-field gel electrophoresis........ 79
2.7.3 Polymerase chain reaction (PCR)........ 80
2.7.4 Agarose gel electrophoresis........ 86
2.7.5 Pulsed field gel electrophoresis........ 86
2.7.6 Polyacrylamide gel electrophoresis (PAGE)........ 87
2.7.7 Gel purification of DNA........ 87
2.7.8 Labeling of DNA probes........ 88
2.7.9 Southern blotting........ 89
2.7.10 Southern and colony hybridization........ 90
2.7.11 DNA sequencing........ 90
2.8 RNA Analysis........ 91

2.8.1 Northern analysis.. 91
2.8.2 High resolution S1 nuclease protection analysis.................... 92
2.8.3 Primer extension analysis.. 93
2.8.4 RT-PCR analysis.. 93
2.9 Protein analysis... 94
2.9.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)............ 94
2.9.2 Western analysis.. 95
2.10 Confocal Microscopy.. 95
2.11 Software and web sources used in analysis............................... 96

CHAPTER 3: RESULTS... 97
3.1 Preparation and analysis of new *S. clavuligerus cea*S2 mutants................. 99
3.1.1 Preparation and analysis of the *cea*S2 frame shift (*cea*S2-Fs) mutant.. 100
3.1.1.1 Preparation of the *cea*S2-Fs mutant........................ 100
3.1.1.2 Phenotype of the *cea*S2-Fs mutant......................... 105
3.1.1.3 Complementation of the *cea*S2-Fs mutant................ 122
3.1.2 Preparation and analysis of the *cea*S2 deletion mutants........... 122
3.1.2.1 Preparation of the Δ*cea*S2*::apr* mutant.................... 123
3.1.2.2 Phenotype of the Δ*cea*S2*::apr* mutant..................... 128
3.1.3 Preparation and analysis of the *cea*S2 in-frame deletion mutant... 135
3.1.3.1 Preparation of the *S. clavuligerus* Δ*cea*S2 mutant........ 135
3.1.3.2 Phenotype of the Δ*cea*S2 mutant............................ 142
3.1.3.3 Complementation of the Δ*cea*S2 mutant.................... 142
3.2 Isolation and DNA sequence of the *cea*S paralogue............................. 149
3.2.1 Isolation and sequencing of the *cea*S paralogue...................... 149
3.2.2 Protein sequence analysis... 155
3.3 Preparation and analysis of the *S. clavuligerus cea*S1 and *cea*S1/*cea*S2 mutants... 165
3.3.1 Preparation of the *cea*S1 mutant.. 170

3.3.2 Phenotype of the Δ*cea*S1*::apr* mutant.................................. 175
3.3.3 Preparation of the *cea*S1/*cea*S2 double mutant........................ 175
3.3.4 Phenotype of the Δ*cea*S1*::apr/cea*S2-Fs double mutant............ 182
3.4 Preparation and analysis of the *c7p* and the *skn* mutants......................... 187
3.4.1 Sequence analysis of *c7p* and *skn*...................................... 187
3.4.2 Preparation of the *c7p* and the *skn* mutants........................... 190
3.4.3 Phenotypes of the Δ*c7p::apr* and the Δ*skn::apr* mutants.......... 191
3.5 Examination of the linkage between the clavulanic acid, the clavam and the paralogue gene clusters.. 196
3.5.1 Chromosome walking studies.. 205
3.5.2 PFGE and Southern analysis... 217
3.6 Regulation of *cea*S1 and *cea*S2 expression... 224
3.6.1 Nutritional regulation of *cea*S1 and *cea*S2 expression.............. 225
3.6.1.1 RT-PCR analysis.. 225
3.6.1.2 Use of e*gfp* for promoter activity analysis................. 228
3.6.2 Effect of CcaR on *cea*S1 and *cea*S2 transcription.................. 231
3.7 Transcriptional mapping of the *cea*S1 and *cea*S2 promoters...................... 234
3.8 Mapping of the *bls*1 and *pah*1 transcripts.. 243

CHAPTER 4: DISCUSSION.. 251
4.1 Isolation of *cea*S1 from the paralogue gene cluster and linkage studies....... 252
4.2 Phenotypes of the *S. clavuligerus cea*S mutants.................................. 256
4.3 Phenotypes of the *S. clavuligerus* Δ*c7p::apr* and Δ*skn::apr* mutants........... 261
4.4 Transcriptional mapping of the early genes from the paralogue gene cluster.. 264
4.5 Nutritional regulation and effect of CcaR on clavam metabolite biosynthesis 265
4.6 Purpose of paralogues genes in clavam metabolite biosynthesis................. 267
4.7 Conclusion and proposed future research... 269

CHAPTER 5: REFERENCES... 272

List of Tables

Page

CHAPTER 1

1.1 Clavam producing bacteria isolated to date... 9

1.2 Clavulanic acid and 5*S* clavam production in mutants defective in genes encoding the enzymes involved in the early stages of clavam metabolite biosynthesis in *S. clavuligerus*.. 30

CHAPTER 2

2.1 *E. coli* strains used in this study... 51

2.2 *S. clavuligerus* strains used in this study.. 52

2.3 Cloning vectors cosmids and recombinant plasmids used in this study......... 54

2.4 Oligonucleotide primers used in this study.. 81

CHAPTER 3

3.1 Clavulanic acid produced by *S. clavuligerus* wild type, *cea*S2*::apr* and *cea*S2-Fs mutant strains after 96 hours of growth on soy medium as determined by HPLC analysis of culture supernatants............................. 109

3.2 Clavaminic acid, clavam-2-carboxylate and 2-hydroxymethyl clavam produced by *S. clavuligerus* wild type, *cea*S2*::apr* and *cea*S2-Fs mutant strains after 96 hours of growth on soy medium as determined by HPLC analysis of culture supernatants... 110

3.3 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type and *cea*S2-Fs mutant strains grown in glycerol containing medium............ 120

3.4 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type and *cea*S2-Fs mutant strains as determined by HPLC analysis of culture supernatants, after 96 hours of growth in SA and soy media supplemented with dehydroacetic acid (DHA)... 121

3.5 Clavulanic acid produced by *S. clavuligerus cea*S1 mutants after 72 and 96 hours of growth on SA and soy media as determined by HPLC analysis of culture supernatants.. 178

3.6 Clavaminic acid, clavam-2-carboxylate and 2-hydroxymethyl clavam produced by *S. clavuligerus cea*S1 mutants after 72 and 96 hours of growth on soy medium as determined by HPLC analysis of culture supernatants...... 179

3.7 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutant strains grown in glycerol containing medium.. 180

3.8 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutant strains as determined by HPLC analysis of culture supernatants, after 96 hours of growth in SA and soy media supplemented with dehydroacetic acid (DHA)........................ 181

3.9 Clavulanic acid produced by *S. clavuligerus* Δ*c7p::apr* mutants after 72 and 96 hours of growth on SA and soy medium as determined by HPLC analysis of culture supernatants.. 199

3.10 Clavulanic acid produced by *S. clavuligerus* Δ*skn::apr* mutants after 72 and 96 hours of growth on SA and soy medium as determined by HPLC analysis of culture supernatants.. 202

3.11 DNA sequence of T7 and T3 ends of cloned inserts from cosmids used in this study.. 211

3.12 Approximate sizes of restriction fragments hybridizing to probes specific for the clavulanic acid, the clavam and the paralogue gene clusters................ 223

CHAPTER 4

4.1 Paralogous or similar genes located in the clavulanic acid, the clavam and the paralogue gene clusters of *S. clavuligerus*...................................... 253

List of Figures

Page

CHAPTER 1

1.2.1 β-lactam metabolites produced by *S. clavuligerus*.............................. 7

1.3.1 The *S. clavuligerus* cephamycin C biosynthetic pathway...................... 13

1.3.2 The *S. clavuligerus* cephamycin C gene cluster.................................. 15

1.4.1 The *S. clavuligerus* clavulanic acid and 5*S* clavam biosynthetic pathway... 22

1.4.2 The *S. clavuligerus* clavulanic acid gene cluster................................ 24

1.4.3 The *S. clavuligerus* clavam gene cluster and the recently isolated *pah*1 gene.. 28

1.4.4 The pathway proposed by Baldwin *et al.* (1994) showing the late steps thought to be involved in clavulanic acid and 5*S* clavam metabolite biosynthesis.. 34

CHAPTER 3

3.1.1 Introduction of the frame shift mutation in *cea*S2.............................. 102

3.1.2 Homologous recombination involving pCAD2-3(L1-5) and the *S. clavuligerus cea*S2*::apr* (4B) mutant chromosome............................. 104

3.1.3 Southern analysis of *cea*S2-Fs mutants.. 107

3.1.4 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, *cea*S2*::apr* (4B) and *cea*S2-Fs mutant strains............................ 112

3.1.5 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type strain as determined by LC-MS analysis....................................... 114

3.1.6 Mass spectra of different clavam compounds analyzed by LC-MS.......... 116

3.1.7 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus cea*S2-Fs (2-8) mutant as determined by LC-MS.. 118

3.1.8 Putative truncated form of CeaS2, arising due to the *cea*S2-Fs mutation... 125

3.1.9 Preparation of the gene disruption cassette using REDIRECT© technology.. 127

3.1.10 Preparation of the mutagenized cosmid 12B8-AP containing the disrupted copy of *cea*S2.. 130

3.1.11 Construction of the *S. clavuligerus* Δ*cea*S2*::apr* mutant using the mutagenized recombinant cosmid 12B8-AP.. 132

3.1.12 Southern analysis of the *S. clavuligerus* Δ*cea*S2*::apr* mutants............... 134

3.1.13 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* Δ*cea*S2*::apr* mutants as determined by HPLC analysis of culture supernatants.. 137

3.1.14 Preparation of pUWL-Δ*cea*S2, the Δ*cea*S2 in-frame disruption construct.. 139

3.1.15 DNA sequence of the region corresponding to the Δ*cea*S2 in-frame deletion in pUWL-Δ*cea*S2.. 141

3.1.16 Schematic of the process leading to the isolation of the Δ*cea*S2 in-frame deletion mutant using the plasmid pUWL-Δ*cea*S2............................... 144

3.1.17 Southern analysis of the *S. clavuligerus* Δ*cea*S2 mutant......................... 146

3.1.18 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* Δ*cea*S2 in-frame deletion mutant.. 148

3.2.1 Isolation of the *cea*S paralogue (*cea*S1)... 152

3.2.2 Analysis of the *cea*S1 coding sequence using FramePlot....................... 154

3.2.3 DNA sequence of *cea*S1.. 157

3.2.4 Diagrammatic representation of the genes from the paralogue gene cluster present on the insert in the recombinant cosmid 14E10.......................... 162

3.2.5 Alignment of CeaS and acetolactate synthase/acetohydroxy synthase (AHAS) proteins from other *Streptomyces*.. 164

3.2.6 CeaS1 and CeaS2 protein alignment.. 167

3.2.7 Predicted tertiary structure of CeaS1 and its comparison to CeaS2.......... 169

3.3.1 Preparation of the Δ*cea*S1*::apr* and the Δ*cea*S1*::apr*/*cea*S2-Fs mutants.... 172

3.3.2 Verification of the Δ*cea*S1*::apr* single mutants by Southern analysis....... 174

3.3.3 Clavulanic acid and 5*S* clavam production by the *S. clavuligerus* Δ*cea*S1*::apr* mutants.. 177

3.3.4 Southern analysis of Δ*cea*S1*::apr/cea*S2-Fs double mutant strains.......... 184

3.3.5 Clavulanic acid and 5*S* clavam production by the *S. clavuligerus* Δ*cea*S1*::apr/cea*S2-Fs mutants... 186

3.3.6 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* Δ*cea*S1*::apr/cea*S2-Fs mutant (2-7-R-1).. 189

3.4.1 Southern analysis of the Δ*c7p::apr* mutants.. 193

3.4.2 Southern analysis of the Δ*skn::apr* mutants... 195

3.4.3 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*c7p::apr* and Δ*skn::apr* mutant strains as determined by HPLC analysis.. 198

3.4.4 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* Δ*c7p::apr* mutant (2-1).. 201

3.4.5 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* Δ*skn::apr* mutant (2-1).. 204

3.5.1 Diagrammatic representation of the cephamycin-clavulanic acid, the paralogue and the clavam gene clusters... 207

3.5.2 Cosmids containing the cephamycin-clavulanic acid and paralogue gene clusters.. 210

3.5.3 Probe preparation for chromosome walking.. 216

3.5.4 Restriction analysis of *pcb*AB containing recombinant cosmid 8-51......... 220

3.5.5 PFGE and Southern analysis of *S. clavuligerus* chromosomal DNA.......... 222

3.6.1 Assessment of *cea*S1 and *cea*S2 transcript levels by RT-PCR................ 227

3.6.2 Preparation of the *cea*S1 and *cea*S2 promoter-*egfp* reporter constructs...... 230

3.6.3 Use of EGFP as a reporter to detect *cea*S1 and *cea*S2 promoter activity.... 233

3.6.4 Clavulanic acid and 5*S* clavams produced by *S. clavuligerus* wild type and Δ*cca*R::*tsr*A mutant strains.. 236

3.7.1 Mapping of the *cea*S1 transcriptional start point........................ 238

3.7.2 Mapping of the *cea*S2 transcriptional start point........................ 240

3.7.3 DNA sequence of the *cea*S1 and *cea*S2 promoter regions.................... 242

3.8.1 Mapping of the polycistronic transcript encoding *bls*1 and *pah*1 using S1 nuclease protection analysis.. 245

3.8.2 Northern analysis of wild type *S. clavuligerus* RNA using a *pah*1 specific probe.. 248

3.8.3 Predicted RNA secondary structure elements in the region downstream of *pah*1.. 250

CHAPTER 4

4.1 Proposed roles of the different regulatory proteins described in this study.. 263

List of Symbols and Abbreviations

A	Adenine/Alanine/Absorbance
α-AAA	α-Aminoadipic acid
Act	Actinorhodin
ACV	δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine
ACVS	ACV synthase
AHAS	Acetohydroxy acid synthase
amp	Ampicillin
amp	Ampicillin resistance marker
apr	Apramycin
apr	Apramycin resistance marker
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BLIP	β-lactamase inhibitory protein
β-LS	β-lactam synthetase
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine/Cysteine/ Carboxy
cam	Chloramphenicol
CAS	Clavaminate synthase
cDNA	Complementary DNA
CeaS	Carboxyethylarginine synthase
ceph-clav	Cephamycin and clavulanic acid
CFE	Cell free extract
Ci	Curie
D	Aspartate
DAC	Deacetylcephalosporin C
DACS	DAC synthase
DAOCS	Deacetoxycephalosporin C synthase
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddNTP	Dideoxynucleoside triphosphate
dGTP	Deoxyguanosine triphosphate
DHA	Dehydroacetic acid
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
dNTP	Deoxynucleoside triphosphate
dTTP	Deoxythymidine triphosphate
E	Glutamate
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
F	Phenylalanine/Farad
FRT	FLP Recombination Target

G	Guanine/Glycine/Free energy
G3P	Glyceraldehyde-3-phosphate
GSPG	Glycerol sucrose proline glutamic acid medium
GTP	Guanosine triphosphate
H	Histidine
HEPES	Hydroxyethylpiperazinethansulfonic acid
HPLC	High performance liquid chromatography
HTH	Helix-turn-helix
I	Isoleucine
IPNS	Isopenicillin N synthase
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISP-4	International *Streptomyces* project medium
K	Lysine
kan	Kanamycin
kan	Kanamycin resistance marker
Kb	Kilobase
kDa	KiloDalton
kV	KiloVolt
L	Leucine
LAT	Lysine ε-aminotransferase
LB	Lennox broth
LC-MS	Liquid chromatography mass spectrometry
M	Methionine/Molar
m	Mass/Meter
Mb	Megabase
MCS	Multiple cloning site
mol	Mole
MOPS	3-[N-Morpholino]propanesulfonic acid
mRNA	Messenger RNA
MYM	Maltose yeast extract malt extract medium
N	Asparagine/Amino
NaTCA	Sodium trichloroacetate
neo	Neomycin resistance marker
OAT	Ornithine acetyltransferase
OD	Optical density
OCT	*o*-carbamoyltransferase
ORF	Open reading frame
oriT	Origin of transfer
P	Proline
P6C	1-piperidine-6-carboxylate
PAGE	Polyacrylamide gel electrophoresis
PAH	Proclavaminate amidinohydrolase
PBP	Penicillin binding protein
P buffer	Protoplast buffer
PCA	Perchloric acid
PCD	Piperidine-6-carboxylate dehydrogenase

PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFGE	Pulsed field gel electrophoresis
PIPES	1,4-Piperazine-bis-(2-ethanosulfonic acid)
pmol	Picomole
ppGpp	Guanosine tetraphosphate
psi	Per square inch
Q	Glutamine
R	Arginine
^{r}	Resistant phenotype
Red	Undecylprodigiosin
RNA	Ribonucleic acid
RNAse	Ribonuclease
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase-PCR
S	Serine
^{s}	Sensitive phenotype
SA	Starch asparagine medium
SDS	Sodium dodecyl sulfate
spec	Spectinomycin
SSC	Standard saline citrate
T	Thiamine/Threonine
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
TES	*N*-tris-[hydroxymethyl]-methyl-2-aminoethane-sulfonic acid
tet	Tetracycline
Tm	Melting temperature
TPP	Thiamine pyrophosphate
Tris	Tris(hydroxylmethyl)-aminomethane
TSB	Trypticase soy broth
TSBM	Trypticase soy broth plus 1% maltose
TSBS	Trypticase soy broth plus 1% starch
TSP	Transcriptional start point
tsr	Thiostrepton
tsr	Thiostrepton resistance marker
V	Valine/Volt/Five
W	Tryptophan
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
Y	Tyrosine
YT	Yeast extract tryptone medium
z	Charge

CHAPTER 1: INTRODUCTION

1. Introduction

1.1 The *Streptomyces*

The genus *Streptomyces* was first proposed by Waksman and Henrici (1943), and comprises Gram-positive, aerobic, chemoheterotrophic, spore forming soil bacteria from the order *Actinomycetales* within the class *Actinobacteria*, which have a DNA G+C content of 69-78 mol % (Korn-Wendisch and Kutzner, 1992; Stackebrandt *et al.*, 1997). They are essentially saprophytic in nature and produce a variety of extracellular enzymes that include amylases, chitinases, cellulases, pectinases, xylanases and proteinases that degrade insoluble organic matter, which is then utilized by the organism to fulfill its energy and nutritional requirements (McCarthy and Williams, 1992). In addition, some of the extracellular enzymes produced by *Streptomyces* spp. are of significant commercial importance, as they are used in industry (McCarthy and Williams, 1992). *Streptomyces* spp. carry out a complex developmental life cycle (Chater and Losick, 1996), and are renowned for their tremendous capacity to produce diverse secondary metabolites, which include approximately 50% of all antibiotics produced by bacteria (Miyadoh, 1993). Many of the secondary metabolites produced by streptomycetes have important applications in medicine as antibacterial, antiviral, antitumor or antifungal agents, whereas some are also used in agriculture as growth promoters, fungicides, antiparasitic agents or herbicides (Korn-Wendisch and Kutzner, 1992). This makes streptomycetes important microorganisms to study, in order to better understand the pathways and mechanisms involved in the biosynthesis of these commercially and industrially important secondary metabolites, so that they can be produced in large amounts, at lower costs.

Streptomyces colonies consist of extensively branched substrate and aerial mycelia, and the colony life cycle begins, when one or more long multinucleoid filaments or hyphae arise from a germinating spore (Hardisson *et al.*, 1978). Growth occurs at hyphal tips with extensive branching (Brana *et al.*, 1982; Gray *et al.*, 1990), giving rise to the vegetative or substrate mycelia, which grow on and into the solid culture medium. This is referred to as the vegetative growth phase, during which the substrate mycelia form extensive networks of hyphae within the culture medium, thereby solubilizing it using extracellular enzymes. As the colony ages and the nutrients in the medium are depleted, there is a cessation in the growth of the substrate mycelia, and branches emerge that grow vertically into the air away from the surface of the culture medium, giving rise to the reproductive aerial mycelia (Miguelez *et al.*, 1994; Wildermuth, 1970).

The aerial mycelia are thought to fulfill their nutritional requirement by cannibalizing the substrate mycelia (Mendez *et al.*, 1985), as degradation of substrate hyphae is observed at specific regions of the colony, with the emergence of the aerial hyphae (Wildermuth, 1970). At this stage, antibiotic production, coinciding with morphological differentiation, is also observed in certain portions of the substrate mycelia (Champness and Chater, 1994), which is thought to ward off competition posed by other microorganisms that may compete for the nutrients released by the degenerating substrate mycelia (Chater and Merrick, 1979). Therefore, the production of secondary metabolites is coordinately regulated with morphological differentiation, and they share some of the same regulatory elements (Champness and Chater, 1994; Chater, 1993).

As the growth cycle progresses, septation or cross wall formation takes place in the aerial hyphae, giving rise to chains of uninucleoid compartments, which later develop into thick-walled unigenomic spores (Hardisson and Manzanal, 1976; Manzanal and Hardisson, 1978). The spores produced are more resistant to environmental conditions than the hyphae, and serve as a means for dispersal, giving rise to new *Streptomyces* colonies elsewhere (Korn-Wendisch and Kutzner, 1992).

Streptomyces chromosomes are linear and approximately 8-9 Mb in size (Bentley *et al.*, 2002; Gravius *et al.*, 1994; Ikeda *et al.*, 2003; Leblond *et al.*, 1993; Lezhava *et al.*, 1995; Lin *et al.*, 1993; Omura *et al.*, 2001) and their ends are thought to somehow associate *in vivo* (Yang and Losick, 2000). In addition, *Streptomyces* also possess linear and circular plasmids (Kinashi, 1994; Schrempf *et al.*, 1975). The linear chromosomes and plasmids from *Streptomyces* contain terminal inverted repeats (TIRs), which vary in size from 174 bp to 550 kb (Bentley *et al.*, 2002; Leblond *et al.*, 1996; Lin *et al.*, 1993; Omura *et al.*, 2001; Pandza *et al.*, 1997). These inverted repeats are covalently bound to terminal proteins (Tpg) at their 5′ ends (Bao and Cohen, 2001; Yang *et al.*, 2002), which in turn interact with three additional proteins that include DNA polymerase I and topoisomerase I (Bao and Cohen, 2004; Bao and Cohen, 2003). *Streptomyces* chromosomes and plasmids replicate bi-directionally (Musialowski *et al.*, 1994) from origins of replications, which are located near the center of the respective linear molecules (Bentley *et al.*, 2002; Calcutt and Schmidt, 1992; Omura *et al.*, 2001; Zakrzewska-Czerwinska and Schrempf, 1992). This mechanism of linear chromosomal replication is thought to generate 3′ telomeric overhangs (Chen, 1996), which are "patched" by the proteins associated with the telomeric ends (Bao and Cohen, 2004).

The complete genomes of *S. coelicolor* A3(2) (Bentley *et al.*, 2002) and *S. avermitilis* ATCC31267 (Ikeda *et al.*, 2003; Omura *et al.*, 2001) have been sequenced. The *S. coelicolor* A3(2) chromosome is 8.66 Mb in size and contains 7,825 predicted open reading frames (ORFs), and 22 gene clusters encoding known or predicted secondary metabolites (Bentley *et al.*, 2002). Similarly, the *S. avermitilis* chromosome is 9.02 Mb in size with 7,574 predicted ORFs and 30 gene clusters, encoding known or predicted secondary metabolites (Ikeda *et al.*, 2003). Interestingly, both the chromosomes of *S. coelicolor* and *S. avermitilis* are biphasic and can be divided into a central core region (approximately 4.9 and 6.5 Mb in *S. coelicolor* and *S. avermitilis*, respectively) containing most of the essential genes, surrounded by the outer arms that contain genes involved in nonessential functions (Bentley *et al.*, 2002; Ikeda *et al.*, 2003). The chromosomes of *S. avermitilis* and *S. coelicolor* also show a very high degree of synteny, especially in the central core regions, and 69% of the predicted *S. avermitilis* ORFs were found to have orthologs in *S. coelicolor* (Ikeda *et al.*, 2003). In addition, the genome sequences show that more than half of the secondary metabolite gene clusters are located in the chromosomal arms and that there are multiple copies of similar (paralogous) genes, which are sometimes arranged in paralogous gene clusters, in both *S. avermitilis* and *S. coelicolor* (Bentley *et al.*, 2002; Ikeda *et al.*, 2003).

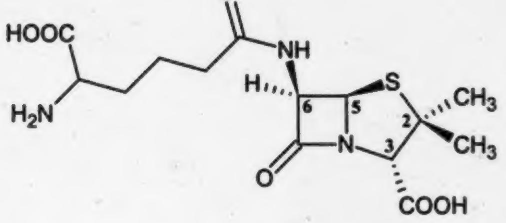
In *Streptomyces*, the genes involved in the biosynthesis of a secondary metabolite are invariably grouped together forming a gene cluster (Chater and Bibb, 1997), which can typically range from 15 to 120 kb in size (Paradkar *et al.*, 2003). This paradigm has proven extremely useful in genetic studies of antibiotic production, where the identification of a single gene provides a 'handle' through which the entire antibiotic gene cluster can be located. Therefore, to date, the genes for approximately 80 actinomycete secondary metabolite pathways have been cloned and sequenced to some extent (Paradkar *et al.*, 2003). Normally in *Streptomyces*, secondary metabolite gene clusters are located on the chromosome, although there are four examples where these gene clusters are present on large plasmids. These are as follows: SCP1 in *S. coelicolor* A3(2) encoding methylenomycin (Kirby *et al.*, 1975; Redenbach *et al.*, 1998), pSLA2-L in *S. rochei* encoding lankacidin and lankamycin (Kinashi *et al.*, 1994; Mochizuki *et al.*, 2003), pPZG103 in *S. rimosus* encoding oxytetracycline (Gravius *et al.*, 1994; Pandza *et al.*, 1998) and pSV1 from *S. violaceoruber* also encoding methylenomycin (Aguilar and Hopwood, 1982; Yamasaki *et al.*, 2003). In addition, antibiotic biosynthesis gene clusters also contain resistance, transport and regulatory genes that are physically linked and are coordinately regulated with the biosynthetic enzymes (Chater and Bibb, 1997).

1.2 The β-Lactam metabolites

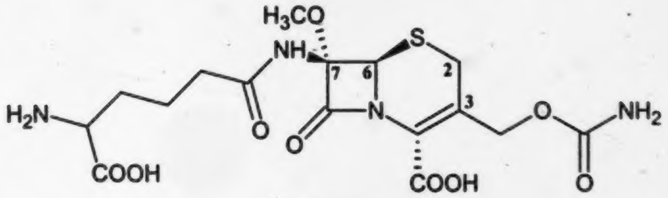
All β-lactam metabolites contain the characteristic four membered azetidinione ring system (also called the β-lactam ring, Figure 1.2.1), which is responsible for their biological activity (Skatrud *et al.*, 1997). The β-lactam antibiotics act by inhibiting the transpeptidation reaction required for peptidoglycan cross-linking, by acting as structural analogues of the acyl-D-alanyl-D-alanine terminus of the stem peptide, which is the substrate of the transpeptidase enzyme (Tipper and Strominger, 1965; Wise and Park, 1965). Inhibition of transpeptidation in growing bacterial cells is thought to weaken the peptidoglycan layers, which provide rigidity and protect the cells from osmotic rupture, thereby eventually leading to cell death by lysis. In addition to transpeptidases, other membrane proteins involved in cell wall biosynthesis/assembly have been also shown to bind penicillin covalently, and are therefore called penicillin-binding proteins (PBPs) (Waxman and Strominger, 1982).

The first β-lactams were accidentally discovered by Alexander Fleming in 1928, when he noticed that a fungal contaminant, later found to be *Penicillium notatum*, could inhibit the growth of staphylococci on culture media (Macfarlane, 1984). The antibacterial material was found to be present in the filtrates from the mould culture (Fleming, 1929), and was later characterized as a mixture of solvent-soluble hydrophobic penicillins, which are only produced by filamentous fungi (Demain and Elander, 1999). Penicillin N and Cephalosporin C, also produced by a fungus, were the first hydrophilic β-lactams to be isolated and characterized (Abraham *et al.*, 1955). In addition to the β-

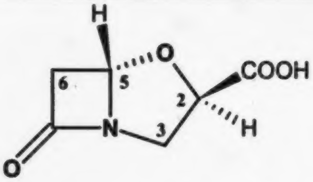
Figure 1.2.1 β-lactam metabolites produced by *S. clavuligerus*. The chemical structures of the different β-lactam metabolites produced by the wild type strain of *S. clavuligerus* are shown. The scheme used to number the residues forming the bicyclic nucleus of penicillin N, cephamycin C, the 5*S* clavams (shown in the hollow box) and clavulanic acid, is also indicated.


HOOC
H_2N
O
NH
H
6
5
S
2
CH_3
CH_3
3
N
O
COOH

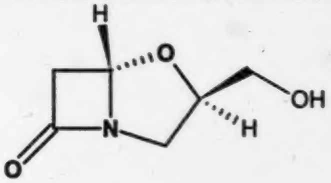
Penicillin N


H_3CO
NH
7
6
S
2
3
H_2N
O
N
O
COOH
COOH
O
NH_2
O

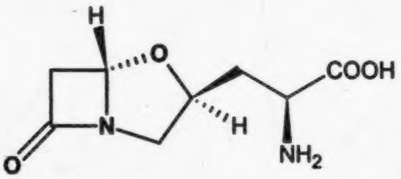
Cephamycin C


H
O
6
5
2
COOH
3
N
H
O

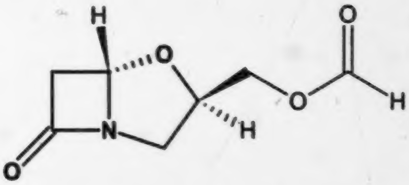
Clavam-2-carboxylate



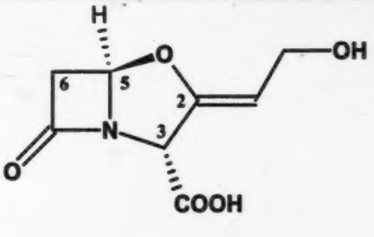
2-Hydroxymethyl clavam


H
O
COOH
N
H
NH_2
O

Alanylclavam



2-Formyloxymethyl clavam


H
O
OH
6
5
2
N
3
O
COOH

Clavulanic acid

lactam ring, the nucleus of the penicillins consists of the characteristic five membered sulfur-containing thiazolidine ring, whereas that of the cephalosporins contains the six membered dihydrothiazine ring (Figure 1.2.1)(Demain and Elander, 1999). The first prokaryotic β-lactam producer ever identified was a penicillin N-producing *Streptomyces* species (Miller *et al.*, 1962), and until then, it was thought that bacteria were not capable of producing any β-lactams (Demain and Elander, 1999). Members of the genus *Streptomyces* and *Nocardia* were also found to produce cephamycins, which are 7-methoxy-cephalosporins that sometimes have side chains attached to the C3 position (Nagarajan *et al.*, 1971; Stapley *et al.*, 1972)(Figure 1.2.1). *Streptomyces clavuligerus*, the subject of this study, produces both cephamycin C and penicillin N in addition to other β-lactam compounds (Higgens and Kastnar, 1971)(Figure 1.2.1).

Several *Streptomyces* species including *S. clavuligerus* also produce clavam metabolites including clavulanic acid (Table 1.1), which are structurally related to each other but differ from the penicillins and cephamycins (Figure 1.2.1). The clavam bicyclic nucleus consists of an oxygen atom-containing oxazolidine ring fused to a β-lactam ring, instead of the sulfur-containing rings that are found in the more conventional penicillin and cephamycin-type antibiotics (Howarth *et al.*, 1976)(Figure 1.2.1). Although clavulanic acid is a very weak β-lactam antibiotic, it is a potent inhibitor of β-lactamases (Brown *et al.*, 1976), which inactivate penicillins and cephalosporins by hydrolyzing their β-lactam rings to form the respective "-oic" acid degradation products (Baggaley *et al.*, 1997). Clavulanic acid acts as a structural analogue of the penicillins and recruits β-lactamases in suicide reactions leading to the formation of stable acyl enzyme intermediates, thereby inactivating the β-lactamases (Massova and Mobashery, 1997). β-lactamases can be further divided into two groups consisting of the active site serine enzymes (class A, C and D) and the zinc-containing metalloenzymes (class B), which have different mechanisms of action (Ambler, 1980). Clavulanic acid inhibits the activity of certain β-lactamases belonging to classes A and D, but it is inactive against the class B metallo-β-lactamases (Walsh, 2003). Due to this β-lactamase inhibitory activity, clavulanic acid is used in combination with other β-lactam antibiotics in commercially available preparations such as Augmentin^{TM} (clavulanic acid + amoxycillin) and

Table 1.1: Clavam producing bacteria isolated to date[a].

Organism	Clavam metabolites	Source/Reference
S. clavuligerus	clavulanic acid, clavam-2-carboxylate, 2-hydroxymethyl clavam, 2-formyloxymethyl clavam and alanylclavam	Brown *et al.* (1976), Brown *et al.* (1979) and Pruess and Kellett (1983)
S. jumonjinensis	clavulanic acid	Box (1977)
S. katsurahamanus	clavulanic acid	Kitano *et al.* (1979)
Streptomyces sp.	clavulanic acid	Sanraku-Ocean Co. (1981)
S. lipmanii	clavaminic acid	Elson *et al.* (1987)
S. antibioticus ssp. *antibioticus* TU1718	2-hydroxyethyl clavam and valclavam	Peter *et al.* (1985) and Wanning *et al.* (1981)
S. hygroscopicus	clavamycins A-F	Naegeli *et al.* (1986)
S. platensis	clavamycin A	Kawamura *et al.* (1987a)
S. lavendulae	clavamycin D, E	Kawamura *et al.* (1987b)
S. brunneogriseus ssp. *bannaensis*	clavamycin E	Seki *et al.* (1987)

[a] Adapted from Jensen and Paradkar (1999)

Timentin™ (clavulanic acid + ticarcillin), to treat infections caused by β-lactamase-producing microorganisms (Jensen and Paradkar, 1999). Out of all the known naturally produced clavam metabolites, only clavulanic acid has been shown to be inhibitory towards β-lactamases (Brown *et al*., 1976), whereas some of the other clavams display antibacterial and antifungal activities (Jensen and Paradkar, 1999). The stereochemistry of clavulanic acid is 3*R*,5*R*, which is different from the 3*S*,5*S* stereochemistry of the other naturally produced clavams (Figure 1.2.1), and the β-lactamase inhibitory activity of clavulanic acid has been attributed to its unique stereochemistry (Baggaley *et al*., 1997). Based on their stereochemistry, clavam metabolites other than clavulanic acid are commonly referred to as the 5*S* clavams. In addition, the 5*S* clavams also lack the carboxyl group that is present at the C3 position in clavulanic acid (Baggaley *et al*., 1997)(Figure 1.2.1).

The β-lactam antibiotics are probably the most widely used of all antibiotics (Demain, 2000). According to estimates, the total world antibiotic market at the dosage form level in 1996 was ~$23 billion (US) per annum (Demain and Elander, 1999), with the β-lactams cornering ~$15 billion (US), or 65% of that market (Elander, 2003). As well, in 1995 the clavulanic acid-containing concoction Augmentin™ alone was the second largest selling antibacterial that year with world sales valued at ~$1.3 billion (US) (Elander, 2003). Since clavulanic acid is produced industrially by fermenting *S. clavuligerus* (Lawrence and Lilly, 1980), the biosynthetic pathways and the regulatory mechanisms controlling the production of clavulanic acid and the 5*S* clavams are a point of great interest, and are the topic of active research in many laboratories.

1.3 Cephamycin C biosynthesis in *S. clavuligerus*

Streptomyces clavuligerus was first isolated in 1971 by Nagarajan *et al*. (1971), during a screen for organisms producing β-lactams with improved resistance to β-lactamases, which included cephamycin C. The cephamycins or 7-methoxy-cephalosporins are formed in a β-lactam biosynthetic pathway, which involves two "early" steps also involved in fungal penicillin biosynthesis, three "intermediate" steps common with fungal cephalosporin C biosynthesis, and the final "late" steps which are involved exclusively in bacterial cephamycin biosynthesis (Paradkar *et al*., 1997)(Figure

1.3.1). As is the case for other antibiotic biosynthetic gene clusters, the *S. clavuligerus* cephamycin biosynthetic gene cluster contains all of the genes involved in the biosynthesis, self-resistance, regulation and transport of cephamycin C (Alexander and Jensen, 1998)(Figure 1.3.2). Interestingly, the gene cluster involved in clavulanic acid biosynthesis resides immediately downstream of the cephamycin gene cluster in *S. clavuligerus* (Aidoo *et al.*, 1993; Ward and Hodgson, 1993).

The biosynthesis of all penicillins and cephalosporins begins with the precursor amino acids cysteine, valine and α-aminoadipic acid (α-AAA, Figure 1.3.1). In fungi, α-AAA is formed as part of the lysine biosynthetic pathway (Bhattarcharjee, 1985), whereas prokaryotes use a different metabolic pathway for lysine biosynthesis that does not utilize α-AAA as an intermediate (Umbarger, 1978). Therefore, bacterial β-lactam producers synthesize α-AAA in a two-step process using lysine as a precursor. In *S. clavuligerus*, lysine is first deaminated by the enzyme lysine ε-aminotransferase (LAT) (Madduri *et al.*, 1989), which is encoded by the *lat* gene (Madduri *et al.*, 1991), to give 1-piperidine-6-carboxylate (P6C, Figure 1.3.1). In the second step, P6C is dehydrogenated by piperidine-6-carboxylate dehydrogenase (PCD)(Fuente *et al.*, 1997) to give α-AAA (Figure 1.3.2), and the gene encoding PCD (*pcd*) was one of the last genes involved in cephamycin C biosynthesis to be identified (Alexander and Jensen, 1998; Perez-Llarena *et al.*, 1998)(Figure 1.3.2).

In *S. clavuligerus*, the enzyme α-aminoadipyl-cysteinyl-valine synthase (ACV synthase) adenylates and activates α-AAA, cysteine and valine, epimerizes L-valine to the D-configuration, and then carries out the non-ribosomal biosynthesis of LLD-aminoadipyl-cysteinyl-valine (ACV) tripeptide (Schwecke *et al.*, 1992)(Figure 1.3.1). The *pcb*AB gene encoding ACV synthase in *S. clavuligerus* has not been completely sequenced, but all of the other *pcb*AB genes isolated and analyzed to date show the presence of three non-ribosomal peptide synthases-like domains or modules, which are specific for α-AAA, cysteine and valine, respectively (Aharonowitz *et al.*, 1993). Next, isopenicillin N synthase (IPNS), also known as ACV cyclase, the product of the *pcb*C gene (Figure 1.3.2), carries out the cyclization of ACV to give isopenicillin N (Jensen *et al.*, 1986), the first β-lactam ring-containing intermediate in the pathway (Figure 1.3.1).

Figure 1.3.1 The *S. clavuligerus* cephamycin biosynthetic pathway. The identities of the intermediates and enzymes involved are shown, and the genes encoding the respective enzymes are also indicated. The carbon atom involved in the stereochemical reaction catalyzed by Isopenicillin N epimerase is shown by the hollow and filled circles in the structures of Isopenicillin N and Penicillin N, respectively. Abbreviations used are as follows: ACV, Aminoadipyl-Cysteinyl-Valine; DAOC, Deacetoxycephalosporin C and DAC, Deacetylcephalosporin C.

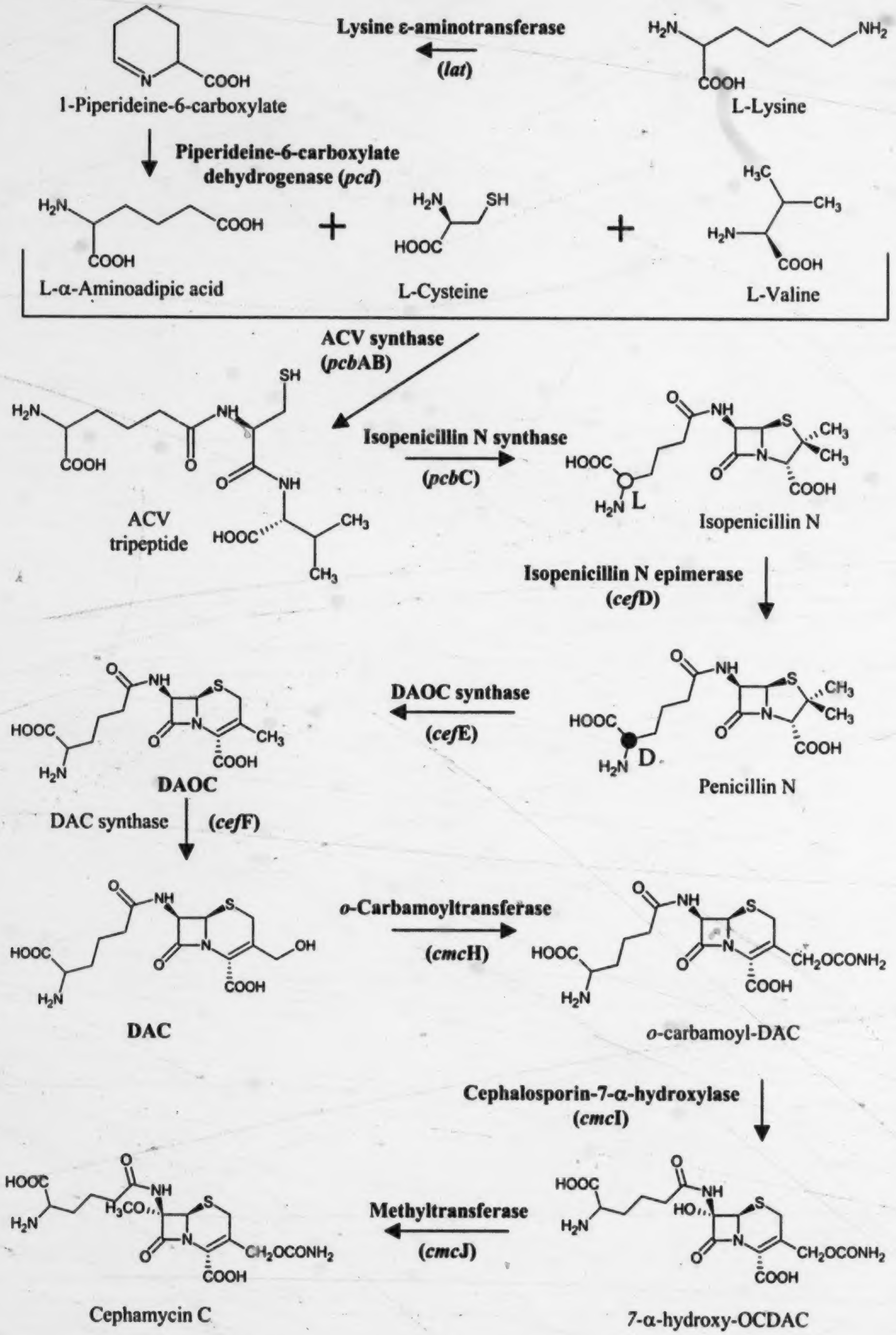
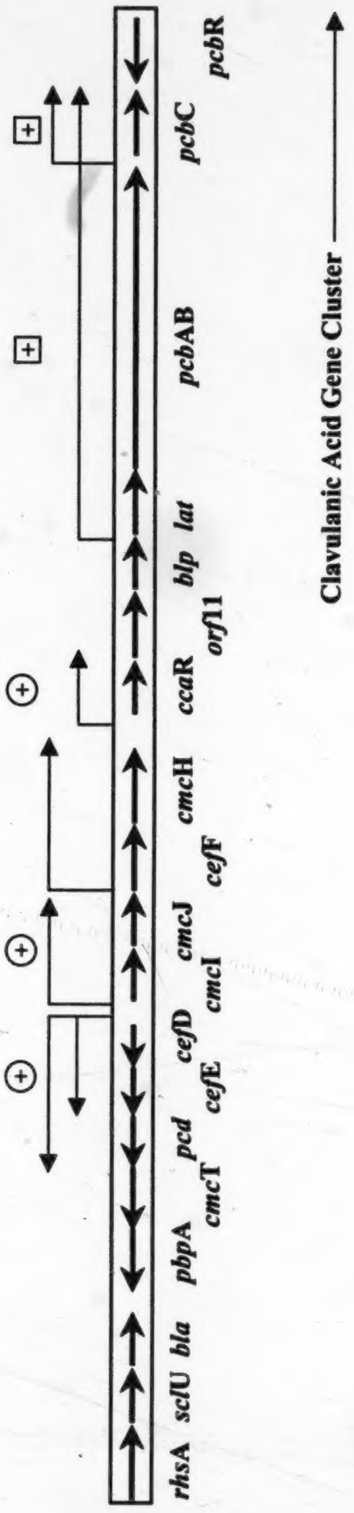

Lysine ε-aminotransferase
(*lat*)
1-Piperideine-6-carboxylate
L-Lysine
Piperideine-6-carboxylate dehydrogenase (*pcd*)
L-α-Aminoadipic acid
L-Cysteine
L-Valine
ACV synthase
(*pcb*AB)
ACV tripeptide
Isopenicillin N synthase
(*pcb*C)
Isopenicillin N
L
Isopenicillin N epimerase
(*cef*D)
Penicillin N
D
DAOC synthase
(*cef*E)
DAOC
DAC synthase
(*cef*F)
DAC
o-Carbamoyltransferase
(*cmc*H)
o-carbamoyl-DAC
Cephalosporin-7-α-hydroxylase
(*cmc*I)
7-α-hydroxy-OCDAC
Methyltransferase
(*cmc*J)
Cephamycin C

Figure 1.3.2 The *S. clavuligerus* cephamycin C gene cluster. The thick arrows represent the respective genes with the arrowheads indicating the direction of transcription. The bent arrows represent the known transcripts that arise from the cephamycin gene cluster. The transcripts that are proposed to be positively regulated by the transcriptional regulator CcaR are indicated by the boxes containing the + sign, whereas the transcripts that are known to be regulated by CcaR are indicated by the circles containing the + sign.


rhsA
sclU
bla
pbpA
cmcT
pcd
cefE
cefD
cmcI
cmcJ
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
pcbC
pcbR
+
+
+
+
+
Clavulanic Acid Gene Cluster

The IPNS enzymes belong to a family of mononuclear ferrous oxygenases that use Fe(II) in catalysis (Roach *et al.*, 1995), and require molecular oxygen, α-ketoglutarate and ascorbate as cofactors (Jensen *et al.*, 1982b).

In penicillin producing filamentous fungi, isopenicillin N is converted into hydrophobic penicillins by side chain substitution (Alvarez *et al.*, 1987), whereas in cephalosporin producing organisms it is converted to penicillin N by the action of the enzyme isopenicillin N epimerase (Usui and Yu, 1989). Isopenicillin N epimerase, encoded by the *cef*D gene (Figure 1.3.2), is a pyridoxalphosphate-requiring racemase that converts the L-α-aminoadipyl sidechain of isopenicillin N to the D configuration (Figure 1.3.1). Next, deacetoxycephalosporin C synthase (DAOCS, also known as penicillin N expandase) that is encoded by the *cef*E gene (Kovacevic *et al.*, 1989)(Figure 1.3.2), carries out the expansion of the five-membered thiazolidine ring of penicillin N, to give the six-membered dihydrothiazine ring found in the cephalosporins (Figure 1.3.1). DAOCS is then hydroxylated to deacetylcephalosporin C (DAC) by the action of deacetylcephalosporin C synthase (DACS, Figure 1.3.1), which is encoded by the *cef*F gene (Kovacevic and Miller, 1991)(Figure 1.3.2). Both DAOCS and DACS are dioxygenases that utilize molecular oxygen, require ferrous ions and α-ketoglutarate, and have many characteristics similar to IPNS (Liras, 1999). Interestingly, prokaryotic DAOCS and DACS are very similar in terms of their amino acid sequences indicating some sort of gene duplication event (Liras *et al.*, 1998), whereas in eukaryotic cephalosporin producers, a single bi-functional protein encodes both DAOCS and DACS activities (Samson *et al.*, 1987).

In fungi, DAC is used for cephalosporin C production, whereas in the actinomycetes it is utilized for the production of cephamycin C by the "late steps" that are specific for cephamycin C biosynthesis (Jensen and Demain, 1995). The first of these steps involves the *o*-carbamoylation at C3 of DAC, carried out by *o*-carbamoyltransferase (OCT), which requires ATP, Mn^{2+}, Mg^{2+} and carbamoylphosphate (Brewer *et al.*, 1980) (Figure 1.3.1). This gives *o*-carbamoyl-desacetycephalosporin C (*o*-carbamoyl-DAC), and the *cmc*H gene encodes OCT in *S. clavuligerus* (Coque *et al.*, 1995)(Figure 1.3.2). In the final steps, the putative *S. clavuligerus* cephalosporin-7-α-hydroxylase (Xiao *et al.*, 1991b) and methyltransferase (Xiao *et al.*, 1991a) enzymes, which are encoded by *cmc*I and *cmc*J, respectively (Alexander and Jensen, 1998)(Figure 1.3.2), carry out the conversion of *o*-carbamoyl-DAC to cephamycin C (Figure 1.3.1).

In addition to genes encoding biosynthetic enzymes, the *S. clavuligerus* cephamycin C gene cluster also contains a regulatory gene, *cca*R (See Section 1.5 for details), and other genes implicated in self-resistance and transport of cephamycin C (Figure 1.3.2). The *cmc*T gene was isolated and described independently by two separate groups (Figure 1.3.2), and is thought to encode a membrane transport protein involved in exporting cephamycin C across the cytoplasmic membrane of *S. clavuligerus* (Alexander and Jensen, 1998; Perez-Llarena *et al.*, 1998). Two genes encoding penicillin-binding proteins (PBPs) are also located at the boundaries of the *S. clavuligerus* cephamycin C gene cluster. The *pcb*R gene (also called *pcb*57) is located at the end of the cephamycin gene cluster that is proximal to the clavulanic acid gene cluster, and *pcb*R mutations are lethal in strains that produce cephamycin C (Paradkar *et al.*, 1996)(Figure 1.3.2). The PcbR protein belongs to the family of high molecular weight, low affinity, class B PBPs, and is thought to be involved in conferring self-resistance to the β-lactam antibiotics produced by *S. clavuligerus* (Paradkar *et al.*, 1996). The second PBP encoding gene (*pbp*A, also called *pbp*74) in the cephamycin gene cluster is found at the end distal to the one encoding PcbR (Alexander and Jensen, 1998)(Figure 1.3.2). The *pbp*A gene is thought to encode a second high molecular weight PBP, which does not display any significant similarity to PcbR, and its function in the cephamycin C biosynthetic gene cluster is still unclear (Perez-Llarena *et al.*, 1998).

In addition to clavulanic acid, *S. clavuligerus* also produces the 17.5-kDa proteinaceous substance called BLIP (for Beta-Lactamase Inhibitory Protein), which is also inhibitory to β-lactamases (Doran *et al.*, 1990a). The *bli* gene that encodes BLIP is not located in the cephamycin gene cluster or in any of the other known gene clusters associated with β-lactam metabolite biosynthesis in *S. clavuligerus* (H-U Park, Personal communication). BLIP inhibits the same types of β-lactamases that are inhibited by clavulanic acid, but it displays a very high degree of variation in its effects on the β-lactamases that it inhibits (Strynadka *et al.*, 1994). There is a second BLIP Like Protein (BLP)-encoding gene (*blp*), which is located in the *S. clavuligerus* cephamycin gene cluster (Perez-Llarena *et al.*, 1997b)(Figure 1.3.2). When a *bli* gene mutant was prepared in a *S. clavuligerus* parental strain that was also unable to produce clavulanic acid, it was shown that this mutant did not produce any β-lactamase inhibitory material, indicating that the *blp* gene product does not have the same bioactivity as BLIP (Thai *et al.*, 2001). In addition, a *blp* mutant did not show any apparent effect on β-lactam antibiotic production (Alexander and Jensen, 1998). Therefore, at this stage the role of BLP in the cephamycin biosynthetic gene cluster in not known.

A β-lactamase-encoding gene (*bla*) is also present in the cephamycin gene cluster and is located at one end of the cluster (Perez-Llarena *et al.*, 1997a)(Figure 1.3.2). The Bla protein is a type A β-lactamase enzyme with low enzymatic activity and no cephalosporinase activity, and is inhibited by both BLIP and clavulanic acid (Perez-Llarena *et al.*, 1997a). Therefore, Bla does not make *S. clavuligerus* self-resistant to cephamycin C, and its exact function remains unclear.

Three ORFs are located in and around the cephamycin C gene cluster for which no functions have been assigned or predicted yet. The putative 338 amino acid gene product encoded by *orf*11 (Figure 1.3.2) does not show any similarity to proteins present in the database (Perez-Llarena *et al.*, 1997b). As well, Alexander *et al.* (1998) demonstrated that an *orf*11 mutant was unaffected in both cephamycin C and clavulanic acid biosynthesis. Interestingly, Southern analysis using an *orf*11-specific probe has demonstrated that it may be conserved in other *Streptomyces* species that also produce clavulanic acid or the 5*S* clavams along with cephalosporins, but not in *Streptomyces* species that only produce cephalosporins (Alexander and Jensen, 1998). The *scl*U and *rhs*A ORFs are thought to reside outside the cephamycin gene cluster, and no function can be predicted for the *scl*U gene product, as it does not show any similarity to any protein present in the database (Alexander and Jensen, 1998)(Figure 1.3.2). The partial DNA sequence of *rhs*A has been determined, and its predicted gene product shows some similarity to putative Rhs proteins from different sources (Alexander and Jensen, 1998). The functions of the Rhs proteins in other organisms are still unclear, but it is known that *rhs* genes are recombination hotspots (Hill, 1999).

The genes involved in cephamycin C biosynthesis are primarily arranged into operons consisting of 2-3 genes that give rise to polycistronic transcripts, thereby the genes present on the same mRNA are transcriptionally co-regulated (See Section 1.5 for details). The "early" genes that include *lat*, *pcb*AB and *pcb*C (encoding LAT, ACVS and IPNS, respectively) are coordinately expressed as a 14 kb polycistronic transcript from the *lat* promoter (Petrich *et al.*, 1994), and the *pcb*C gene is also expressed from its own promoter as a 1.2 kb monocistronic transcript (Petrich *et al.*, 1992)(Figure 1.3.2). The other "early" gene (*pcd*) is clustered, and also transcribed with the *cef*D and *cef*E genes involved in the "intermediate" steps (Figure 1.3.1), as part of a 4.1 kb polycistronic transcript; in addition, the *cef*D and *cef*E genes are also transcribed as a 2.6 kb bicistronic transcript from the *cef*D promoter (Perez-Llarena *et al.*, 1998)(Figure 1.3.2). In another study, northern analysis indicated that a transcript of more than 10 kb in size also comprised *cef*D and *cef*E (Kovacevic *et al.*, 1990). Therefore, it is plausible that this 10 kb transcript may contain the ORFs starting with *cef*D all the way down to *pbp*A (Figure 1.3.2).

The "late" genes involved specifically in the biosynthesis of cephamycin C are also organized into operons. The *cef*F and *cmc*H genes (encoding DACS and OCT respectively, Figure 1.3.1) are transcribed together and on mRNAs that are 2.8 and 4.1 kb in size (Liras, 1999)(Figure 1.3.2). The *cmc*I and *cmc*J genes that encode enzymes involved in the final stages of cephamycin C biosynthesis (Figure 1.3.1), are also co-transcribed as a 1.6 kb mRNA (Liras, 1999)(Figure 1.3.2).

The transcriptional regulator *cca*R (See Section 1.5 for details) and the *blp* gene from the cephamycin gene cluster are transcribed as 0.9 kb (Perez-Llarena *et al.*, 1997b) and 1.2 kb (Liras, 1999) monocistronic mRNAs, respectively (Figure 1.3.2). In addition, recent studies have indicated that the *cca*R gene may be transcribed on two different mRNAs that are ~1.1 kb and ~1.5 kb in size (Bignell *et al.*, 2004), which originate from two separate promoters (Wang *et al.*, 2004).

Overall, the genes and enzymes involved in cephamycin C biosynthesis in *S. clavuligerus* have been subjected to many studies and are well characterized, in terms of both their function and their regulation.

1.4 Biosynthesis of Clavulanic acid and the 5*S* clavams

Although the biosynthetic pathway and the genes involved in cephamycin C production in *S. clavuligerus* have been well studied and a lot is known about them, the corresponding information about the biosynthesis of clavulanic and the 5*S* clavams has only started to emerge in the last ten years or so. Even though separate pathways produce cephamycin C and clavulanic acid (Figure 1.4.1), the two gene clusters involved in their biosynthesis are found grouped together on the *S. clavuligerus* chromosome, forming a β-lactam supercluster (Aidoo *et al.*, 1994; Ward and Hodgson, 1993)(Figure 1.4.2).

In the case of clavulanic acid and the 5*S* clavams, it is known that they share a common biosynthetic pathway at least up to the level of the intermediate, clavaminic acid (Egan *et al.*, 1997)(Figure 1.4.1). A debate has existed over the last 30 years or so about the identities of the precursors that are used in clavulanic acid and 5*S* clavam metabolite biosynthesis. The search for these precursors mostly involved feeding studies using different ^{14}C labeled compounds, and following their incorporation into clavulanic acid. It was initially shown that clavulanic acid was formed by the condensation of two molecules called the C_3 and the C_5 subunits that consisted of 3 and 5 carbon atoms, respectively (Elson, 1979; Elson and Oliver, 1978). The earliest studies suggested that the urea amino acids arginine or ornithine were the precursors for the C_5 subunit (Romero *et al.*, 1986; Townsend and Ho, 1985a). Later it was shown that arginine was the direct precursor, giving rise to the C_5 subunit (Valentine *et al.*, 1993). The identity of the C_3 precursor has remained elusive until recently, and initial feeding studies showed that glycerol, glycerate, propionate and β-hydroxypropionate, could all be incorporated into clavulanic acid (Elson and Oliver, 1978; Gutman *et al.*, 1985; Townsend and Ho, 1985b). In 1999, the enzyme catalyzing the first reaction in clavulanic acid and 5*S* clavam biosynthesis was characterized, and the C_3 subunit was shown to arise from glyceraldehyde-3-phosphate (G3P)(Khaleeli *et al.*, 1999).

The biosynthesis of clavulanic acid and the 5*S* clavams begins with the condensation of a molecule of G3P and L-arginine (Khaleeli *et al.*, 1999), to give N^2-(2-carboxyethyl)arginine (Elson *et al.*, 1993b), the first dedicated intermediate in the biosynthesis of clavulanic acid and the 5*S* clavams (Elson *et al.*, 1993a)(Figure 1.4.1). This first step is catalyzed by the enzyme carboxyethylarginine synthase (CeaS) in a reaction that requires thiamine pyrophosphate (TPP) (Khaleeli *et al.*, 1999). The crystal structure of CeaS was solved recently and the holoenzyme was shown to be a tetramer consisting of two tightly bound dimers, and its quaternary structure was found to be very

Figure 1.4.1 The *S. clavuligerus* clavulanic acid and 5*S* clavam biosynthetic pathway. The identities of the intermediates and enzymes involved are shown. The paralogous pairs of genes that were isolated as part of this and additional studies, which encode the respective enzymes, are also shown. The latter part of the pathway involved in the biosynthesis of clavulanic acid and the 5*S* clavams, which is still not known, is also indicated.

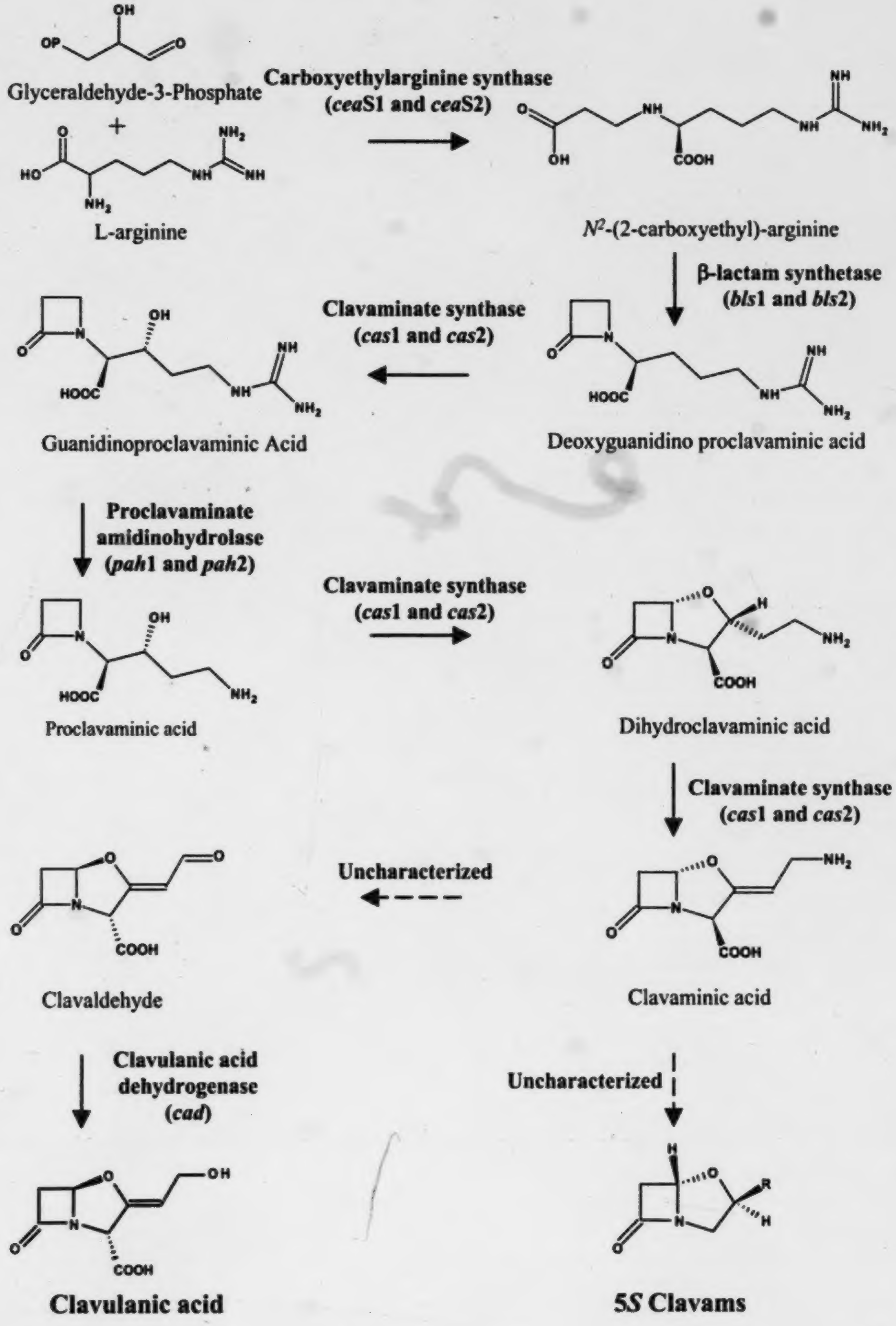
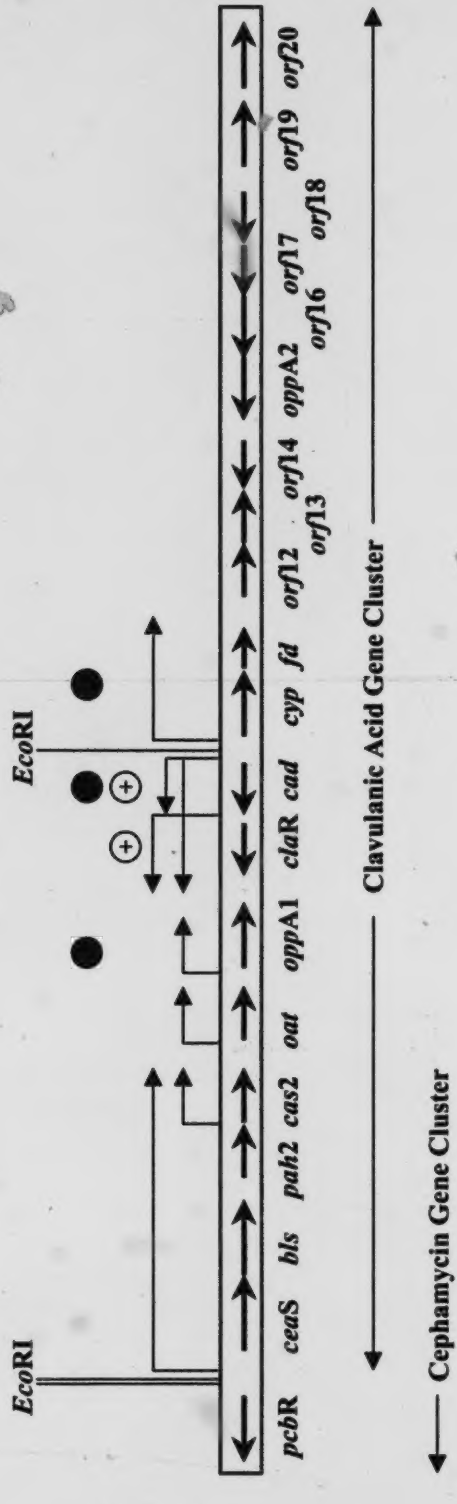

Glyceraldehyde-3-Phosphate
+
L-arginine
Carboxyethylarginine synthase
(*cea*S1 and *cea*S2)
N^2-(2-carboxyethyl)-arginine
β-lactam synthetase
(*bls*1 and *bls*2)
Deoxyguanidino proclavaminic acid
Clavaminate synthase
(*cas*1 and *cas*2)
Guanidinoproclavaminic Acid
Proclavaminate amidinohydrolase
(*pah*1 and *pah*2)
Proclavaminic acid
Clavaminate synthase
(*cas*1 and *cas*2)
Dihydroclavaminic acid
Clavaminate synthase
(*cas*1 and *cas*2)
Clavaminic acid
Uncharacterized
Clavaldehyde
Clavulanic acid dehydrogenase
(*cad*)
Clavulanic acid
Uncharacterized
5*S* Clavams

Figure 1.4.2 The *S. clavuligerus* clavulanic acid gene cluster. The thick arrows represent the respective genes with the arrowheads indicating the direction of transcription. The bent arrows represent the known transcripts that arise from the clavulanic acid gene cluster. The transcripts that are positively regulated by CcaR are indicated by the hollow circles containing the + sign and the transcripts that are positively regulated by ClaR are indicated by the filled circles. Only relevant restriction sites described in the text are shown.


EcoRI
EcoRI
pcbR
ceaS
bls
pah2
cas2
oat
oppA1
claR
cad
cyp
fd
orf12
orf13
orf14
oppA2
orf16
orf17
orf18
orf19
orf20
Cephamycin Gene Cluster
Clavulanic Acid Gene Cluster

similar to the yeast pyruvate decarboxylase (PDC)(Caines *et al.*, 2004), which is involved in the biosynthesis of branched chain amino acids.

Next, the enzyme β-lactam synthetase (β-LS) catalyzes the intramolecular ring closure of N^2-(2-carboxyethyl)arginine to form the β-lactam ring-containing intermediate, deoxyguanidinoproclavaminate (Bachmann *et al.*, 1998; McNaughton *et al.*, 1998) (Figure 1.4.1). The reaction catalyzed by β-LS requires ATP and Mg^{2+}, and the crystal structure of β-LS has been solved showing that the holoenzyme consists of a dimer (Miller *et al.*, 2001). In addition, β-LS shows some similarity to asparagine synthases (AS-B), which belong to a family of enzymes called Ntn amidotransferases (Brannigan *et al.*, 1995).

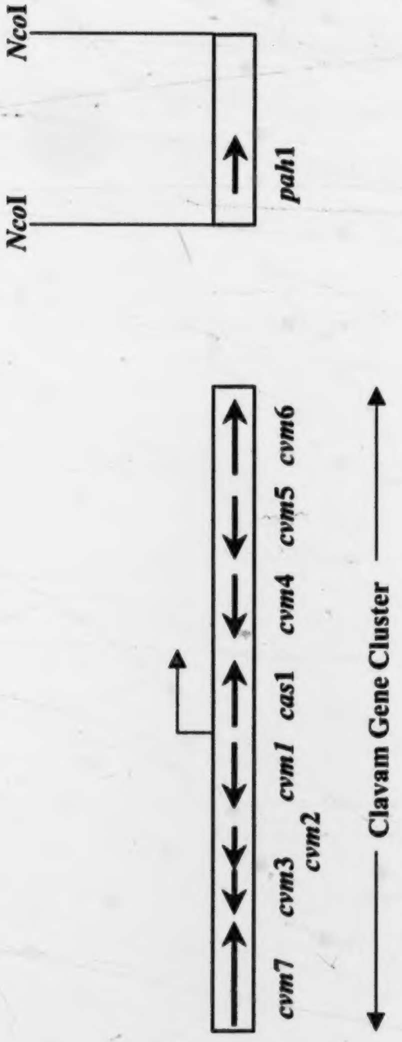
Subsequently, deoxyguanidinoproclavaminate is hydroxylated in the first of three reactions catalyzed by the enzyme clavaminate synthase (CAS), to form guanidinoproclavaminate (Baldwin *et al.*, 1993)(Figure 1.4.1). Clavaminate synthase is a nonheme iron containing, 2-ketoglutarate dioxygenase, which uses molecular oxygen and introduces the hydroxyl group present in guanidinoproclavaminate (Salowe *et al.*, 1990). In addition, CAS belongs to the same family of enzymes as IPNS and DAOCS that are involved in cephamycin C biosynthesis (Figure 1.3.1), even though its amino acid sequence does not show any significant similarity to the respective proteins from the cephamycin gene cluster (Figure 1.3.2). Guanidinoproclavaminate is then converted to proclavaminate by the removal of the guanidino group in a reaction catalyzed by the enzyme proclavaminate amidinohydrolase (PAH)(Elson *et al.*, 1993a)(Figure 1.4.1). The PAH holoenzyme has been shown to consist of a hexamer in solution and requires Mn^{2+} ions for activity; in addition it shows considerable similarity to primary metabolic arginases (Elkins *et al.*, 2002).

Next, CAS mediates the formation of the bicyclic nucleus of clavulanic acid and the 5*S* clavams in a two-step reaction involving oxidative cyclization giving dihydroclavaminic acid (Baldwin *et al.*, 1991), followed by desaturation to form clavaminate (Salowe *et al.*, 1991)(Figure 1.4.1). Clavaminate is thought to be the branch point between the biosynthetic pathways leading to clavulanic acid and the 5*S* clavams (Egan *et al.*, 1997)(Figure 1.4.1). The pathway beyond clavaminate is not well characterized and the only other known intermediate between clavaminate and clavulanic acid is clavaldehyde (Nicholson *et al.*, 1994). Clavaldehyde has the same stereochemistry as clavulanic acid and shows β-lactamase inhibitory activity, and is reduced to clavulanic acid by the action of the NADPH dependent enzyme, clavulanic acid dehydrogenase (CAD)(Nicholson *et al.*, 1994)(Figure 1.4.1). However, the mechanism by which clavaminate undergoes stereochemical inversion and side chain modification to form clavaldehyde is unknown. Similarly, the reactions leading from clavaminate to the 5*S* clavams are not known.

Studies conducted by growing *S. clavuligerus* cultures in an atmosphere of $^{18}O_2$ have demonstrated that the ^{18}O label was incorporated with equal efficiency into the oxazolidine ring and the hydroxyl group present in clavulanic acid (Townsend and Krol, 1988). Since CAS is responsible for the insertion of the oxygen atom found in the oxazolidine ring of clavulanic acid, it has been suggested that the conversion of clavaminic acid into clavaldehyde involves oxidative deamination using molecular oxygen, but the oxygenases(s) or intermediates involved have not been identified yet (Baggaley *et al.*, 1997).

The CAS protein is probably the best-characterized enzyme from the clavulanic acid and 5*S* clavam biosynthetic pathways. When CAS was first purified from *S. clavuligerus*, two isozymic forms of the enzyme were observed, which differed slightly in their kinetic properties and molecular weights (Salowe *et al.*, 1990). Later, it was shown that there were two genes encoding the respective CAS isozymes in *S. clavuligerus*, and their predicted gene products shared 82% identity and 87% similarity, at the amino acid level (Marsh *et al.*, 1992). The two genes encoding CAS were named *cas*1 and *cas*2, and *cas*2 was found to be located downstream of the cephamycin C gene cluster in what is now called the clavulanic acid gene cluster (Aidoo *et al.*, 1994; Ward and Hodgson, 1993)(Figure 1.4.2). Mosher *et al.* (1999) showed that the *cas*1 gene was not located in the clavulanic acid gene cluster, but instead it formed a separate gene cluster along with other genes, which seem to be involved in 5*S* clavam biosynthesis only, and not in the biosynthesis of clavulanic acid (Figure 1.4.3). The genes encoding CAS have been subjected to extensive mutational and transcriptional studies and the *cas*2 gene mutant was first prepared and analyzed by Paradkar and Jensen (1995). It was found that the *cas*2 mutant could still produce clavulanic acid and the 5*S* clavams when grown in

Figure 1.4.3 The *S. clavuligerus* clavam gene cluster and the recently isolated *pah*1 gene. The thick arrows represent the respective genes with the arrowheads indicating the direction of transcription. The bent arrows represent the known transcript that arises from the clavam gene cluster. Only relevant restriction sites described in the text are shown.


NcoI
NcoI
pah1
cvm7
cvm3
cvm2
cvm1
cas1
cvm4
cvm5
cvm6
Clavam Gene Cluster

complex soy medium, but not in defined starch asparagine (SA) medium (Table 1.2). In comparison the wild type strain produced the respective metabolites in both media tested (Table 1.2). This ability of the *cas*2 mutant to produce both clavulanic acid and the 5*S* clavams in soy medium, albeit at reduced levels as compared to the wild type strain, was attributed to the presence of the *cas*1 gene, which encodes the CAS1 isozyme in *S. clavuligerus*. Transcriptional analysis showed that *cas*1 was expressed in soy medium only and not in SA medium, whereas *cas*2 was transcribed in both media tested (Paradkar and Jensen, 1995). Therefore, CAS1 was available for clavulanic acid and 5*S* clavam biosynthesis, only when the *cas*2 mutant was grown in soy medium. This hypothesis was further verified when a *cas*1 gene mutant was prepared and tested by Mosher *et al.* (1999), who showed that the *cas*1 mutant produced clavulanic acid and the 5*S* clavams in soy media, and only clavulanic acid in SA media (Table 1.2), as CAS2 was present for functional replacement of CAS1 in both media tested. In addition, the *cas*1/*cas*2 double mutant was completely blocked in the biosynthesis of all clavam compounds in both soy and SA media (Mosher *et al.*, 1999).

Since its initial isolation, the sequence of the clavulanic acid gene cluster has been extending gradually, and as of now the published sequence is close to ~27 kb and it comprises 19 ORFs (Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Li *et al.*, 2000; Mellado *et al.*, 2002). The first gene in the cluster is *cea*S (Figure 1.4.2), which encodes the CeaS enzyme catalyzing the first reaction in the pathway (Perez-Redondo *et al.*, 1999)(Figure 1.4.1), followed by the *bls* gene, encoding β-LS (Bachmann *et al.*, 1998; McNaughton *et al.*, 1998). Next is *pah* encoding the PAH enzyme (Wu *et al.*, 1995)(Figure 1.4.1). After the two paralogous *cas* genes, *pah* was the third gene involved in clavulanic acid to be subcloned and sequenced (Aidoo *et al.*, 1994), and only later was it shown that *cas*2 was located next to *pah* (Ward and Hodgson, 1993)(Figure 1.4.2). Therefore, except for *cas*1, all of the genes involved in the early part of the pathway that is shared between clavulanic acid and the 5*S* clavams are clustered together and arranged in the same orientation (Figure 1.4.2).

Interestingly, the *oat* gene immediately downstream of *cas*2 encodes an active ornithine acetyltransferase enzyme (OAT) (Elkins *et al.*, 2004; Kershaw *et al.*, 2002)(Figure 1.4.2). OATs are normally involved in arginine biosynthesis and catalyze

Table 1.2: Clavulanic acid and 5*S* clavam production in mutants defective in genes encoding the enzymes involved in the early stages of clavam metabolite biosynthesis in *S. clavuligerus* [a].

Mutant Strain	Respective metabolites produced as compared to the wild type strain in soy medium[b]				
	Clavulanic acid		5*S* clavams[c]		Reference
	Soy	SA	Soy	SA	
Wild type	100	20	100	5-10	Jensen *et al.* (2000)
*cea*S	2-5	0	2-5	0	Jensen *et al.* (2000)
bls	10	0	15	0	Jensen *et al.* (2000)
*pah*2	60	0	100	ND[d]	Jensen *et al.* (2000)
*pah*1	33-41	4-6	ND[d]	0	Jensen *et al.* (2004b)
*pah*1/ *pah*2	0	0	0	0	Jensen *et al.* (2004b)
*cas*2	60	0	60	0	Jensen *et al.* (2000)
*cas*1	27-71	~20	28-87	0	Mosher *et al.* (1999)
*cas*1/ *cas*2	0	0	0	0	Mosher *et al.* (1999)
oat	40	0	2-5[e]	0	Jensen *et al.* (2000)

[a] Adapted from Jensen *et al.* (2000)

[b] The amount of the respective metabolite produced by the wild type strain of *S. clavuligerus* in soy medium was taken to be equal to 100%

[c] The 5*S* clavam metabolites are grouped together and are shown as a family of related compounds

[d] Not determined, the amount of the respective metabolites produced by the strain in question were not determined, or too much variation in the production levels was observed to make any deductions

[e] None of the clavams were detected by HPLC analysis of culture supernatants but some alanyl clavam bioactivity was detected

the step in the arginine biosynthetic pathway involving the transfer of an acetyl group from *N*-acetylornithine to glutamate, thereby recycling the acetyl group and producing ornithine (Cunin *et al.*, 1986). Since arginine is a precursor of clavulanic acid and the 5*S* clavams, it is thought that the function of OAT may be to increase arginine production and channel it towards the biosynthesis of the clavam metabolites, although there is no clear evidence supporting this hypothesis.

When *S. clavuligerus* strains defective in *cea*S, *bls*, *pah* and *oat* genes were prepared and tested (Jensen *et al.*, 2000), their phenotypes were very similar to that of the *cas*2 mutant described earlier (Paradkar and Jensen, 1995)(Table 1.2). All of the gene mutants mentioned above still retained the ability to produce some clavulanic acid and 5*S* clavams in soy, but not in SA medium. Due to the similarities in the observed phenotypes of the *cea*S, *bls*, *pah* and *oat* mutants and the *cas*2 mutant, Jensen *et al.* (2000) proposed that second, paralogous copies of *cea*S, *bls*, *pah* and *oat* exist in *S. clavuligerus*, which are nutritionally regulated in a manner similar to *cas* paralogues. Therefore, it was postulated that the second, paralogous copies of the respective genes are only expressed in soy medium (discussed in detail later).

The only other genes from the clavulanic acid gene cluster with well-defined roles in clavam metabolite biosynthesis are *cla*R and *cad*, which encode the pathway-specific transcriptional regulator for clavulanic acid biosynthesis (See Section 1.5 for details) and CAD (Figure 1.4.1), respectively (Paradkar *et al.*, 1998; Perez-Redondo *et al.*, 1998).

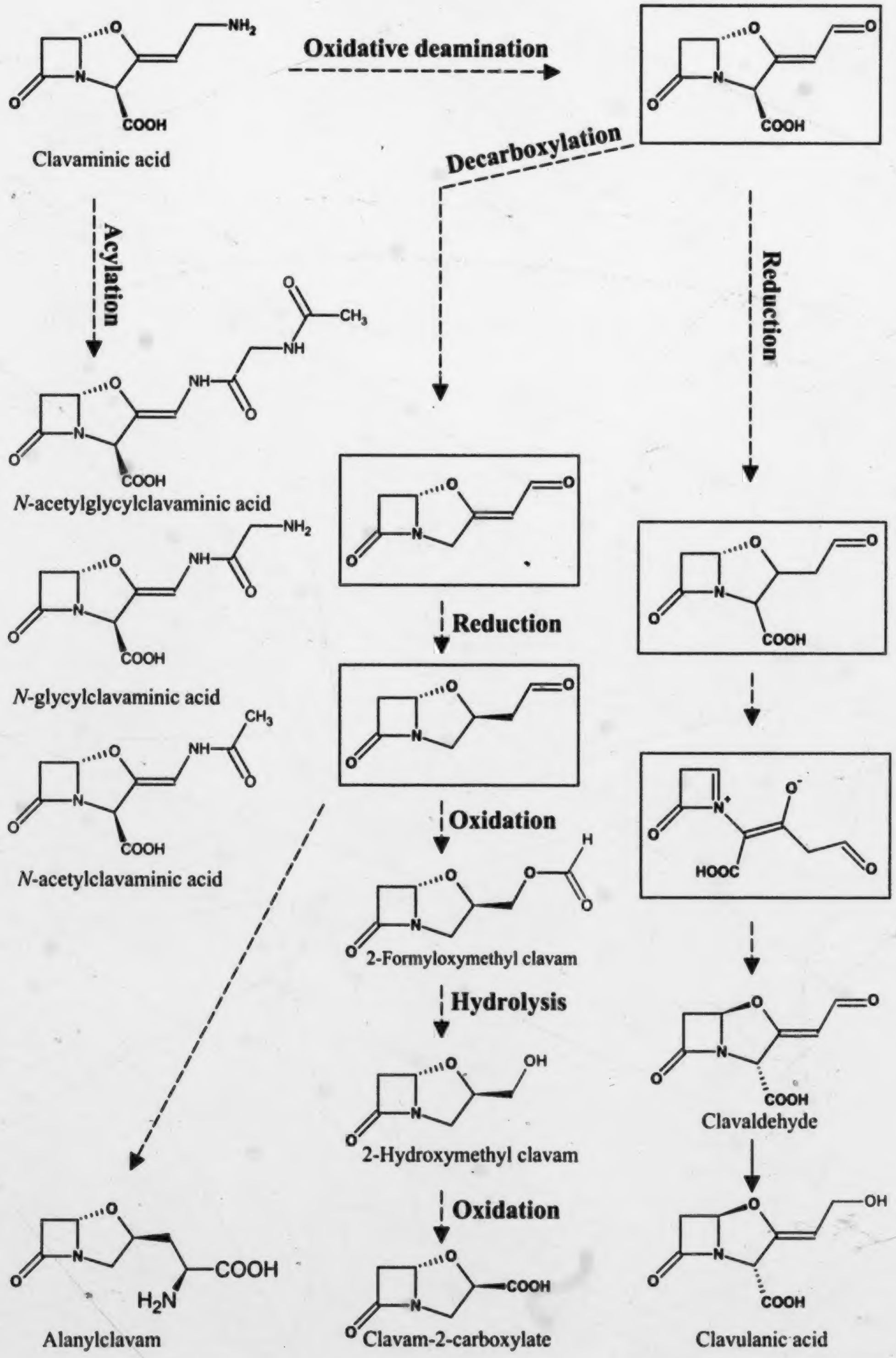
The *opp*A1 and *opp*A2 genes (formerly known as *orf7* and *orf15*, respectively) from the clavulanic acid gene cluster encode putative proteins that show some similarity to oligopeptide binding proteins (Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Mellado *et al.*, 2002)(Figure 1.4.2). Mutants defective in either gene are completely blocked in clavulanic acid biosynthesis in both soy and SA media, but still produced the 5*S* clavams in soy media, demonstrating that they are somehow involved in clavulanic acid biosynthesis (Jensen *et al.*, 2000; Jensen *et al.*, 2004a). Recently it was also proposed that both the *opp*A1 and *opp*A2 genes encode proteins that are responsible for importing short peptides, which can then function as sources of nitrogen (Lorenzana *et al.*, 2004). The ORF immediately downstream of *opp*A2, referred to as *orf16* (Figure 1.4.2), does not show any significant similarity to any known or predicted protein (Jensen *et al.*, 2004a;

Mellado *et al.*, 2002). In the study conducted by Jensen *et al.* (2004a), it was shown that the *orf*16 gene mutant was unable to produce any detectable levels of clavulanic acid, whereas 5*S* clavam production was unaffected. In addition, they also found that both the *opp*A2 and the *orf*16 mutants were accumulating novel 5*S* clavam metabolites, which were acylated derivatives of clavaminic acid. These acylated 5*S* clavam metabolites are thought to represent metabolites from the portion of the clavulanic acid pathway leading from clavaminic acid to clavaldehyde (Elson *et al.*, 1988)(Figures 1.4.4), providing further evidence for the involvement of *opp*A2 and *orf*16 in clavulanic acid biosynthesis.

Two genes in the clavulanic acid gene cluster encode predicted proteins that are similar to cytochrome P-450 monooxygenases from different sources. The first of these is called *cyp* and is located next the *fd* gene that encodes a putative ferredoxin protein (Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Li *et al.*, 2000; Mellado *et al.*, 2002) (Figure 1.4.2), which are electron transport proteins capable of forming complexes with cytochrome P-450 enzymes (O'Keefe and Harder, 1991). When a *S. clavuligerus cyp* mutant was prepared, it was completely blocked in clavulanic acid biosynthesis, whereas it still produced some of the 5*S* clavams in soy medium (Jensen *et al.*, 2000). When a similar *fd* gene mutant was prepared, it still produced clavulanic acid and the 5*S* clavams in soy media, and only clavulanic acid in SA media at roughly 30% of the level as compared to the wild type strain (Jensen *et al.*, 2004a). It has been proposed that *cyp* and *fd* may encode the missing oxygenase activity required for the oxidative deamination of clavaminic acid, resulting in clavaldehyde (Figure 1.4.1), although this has not demonstrated experimentally. The other cytochrome P-450 gene, called *orf*20, has also been knocked out with no apparent effect on clavulanic acid or 5*S* clavam metabolite biosynthesis, and *orf*20 was the last gene from the clavulanic acid gene cluster region of the *S. clavuligerus* chromosome to be sequenced (Jensen *et al.*, unpublished) (Figure 1.4.2).

Three additional ORFs are located between *fd* and *opp*A2 (Figure 1.4.2). They are *orf*12, *orf*13 and *orf*14, encoding a putative β-lactamase, an export/efflux protein and an acetyltransferase, respectively (Jensen *et al.*, 2004a; Mellado *et al.*, 2002). Although the protein encoded by *orf*12 does show some similarity to β-lactamases, it is missing some of the key residues required for activity (Li *et al.*, 2000). In addition, the *orf*12 gene

Figure 1.4.4 The pathway proposed by Baldwin *et al.* (1994) showing the late steps thought to be involved in clavulanic acid and 5*S* clavam metabolite biosynthesis. The known reactions are shown by continuous arrows, whereas the hypothesized reactions are indicated by dashed arrows and the proposed intermediates are boxed. The nature of some of the reactions thought to be involved in the pathway are indicated and the *N*-acylated clavaminic acid derivatives isolated by Elson *et al.* (1998) and Jensen *et al.* (2004) are also shown.


Clavaminic acid
Oxidative deamination
Decarboxylation
Acylation
Reduction
N-acetylglycylclavaminic acid
N-glycylclavaminic acid
N-acetylclavaminic acid
Reduction
Oxidation
2-Formyloxymethyl clavam
Hydrolysis
2-Hydroxymethyl clavam
Oxidation
Alanylclavam
Clavam-2-carboxylate
Clavaldehyde
Clavulanic acid

mutant is completely blocked in clavulanic acid biosynthesis, indicating that it has some other function and it probably does not encode an active β-lactamase (Li *et al.*, 2000) (Jensen *et al.*, 2004a). When the *orf*13 and *orf*14 mutants were prepared and analyzed, they were severely compromised in their ability to produce clavulanic acid, also implicating them in clavulanic acid biosynthesis (Jensen *et al.*, 2004a). As well, it has been suggested that the *orf*13 gene product may be involved in the export of clavulanic acid and/or other clavam metabolites across the *S. clavuligerus* cell membrane.

Based on sequence analysis, the *orf*17 gene product from the clavulanic acid gene cluster is predicted to function as a carboxylase (Jensen *et al.*, 2004a; Mellado *et al.*, 2002). Jensen *et al.* (2004a) prepared an *orf*17 mutant and showed that it was completely blocked in clavulanic acid biosynthesis, indicating that it is involved in a still uncharacterized step in the clavulanic acid biosynthetic pathway.

Genes for two predicted PBPs (*orf*18 and *orf*19, Figure 1.4.2) are also present at one end of the known clavulanic acid gene cluster and are arranged next to each other (Jensen *et al.*, 2004a; Mellado *et al.*, 2002). Mutagenesis studies involving *orf*18 have indicated that it may be an essential gene, and an *orf*19 mutant does not seem to have any effect on clavulanic acid biosynthesis (Jensen *et al.*, 2004a). Therefore, it is unclear at this stage if both of these genes (along with *orf*20) actually reside within the boundaries of the clavulanic acid gene cluster or if they have some other function in *S. clavuligerus*.

As mentioned earlier the *cas*1 gene, encoding the clavaminate synthase isozyme 1, is not located in the clavulanic acid gene cluster. Mosher *et al.* (1999) sequenced the region of the *S. clavuligerus* chromosome surrounding *cas*1 and isolated six new ORFs, of which *cvm*3 and *cvm*6 were only partially sequenced (Figure 1.4.3). Since then, Jensen *et al.* (unpublished) have obtained the complete DNA sequence of *cvm*3 and *cvm*6 and have isolated a number of additional ORFs including *cvm*7, encoding a putative transcriptional regulator similar to the recently isolated pimaricin regulator (PimR), which is involved in the positive regulation of pimaricin biosynthesis in *S. natalensis* (Anton *et al.*, 2004)(Figure 1.4.3). The predicted gene products encoded by *cvm*1, *cvm*2 and *cvm*3 (Figure 1.4.3) show similarities to aldo-keto reductases, ribulose-5-phosphate epimerases and putative oxidoreductases, respectively. As well, the predicted amino acid sequences of *cvm*4, *cvm*5 and *cvm*6 show similarities to acetyltransferases, luciferases and aminotransferases, respectively (Figure 1.4.3). When a *cvm*1 gene mutant was prepared it was shown to be blocked in 5*S* clavam production, but it still produced clavulanic acid in both soy and SA media (Mosher *et al.*, 1999). In the same study, a *cvm*4-*cvm*5 double mutant strain was also prepared and was shown to be blocked in 5*S* clavam, but not clavulanic acid biosynthesis. Therefore, the gene cluster encompassing *cas*1 and the surrounding ORFs has since been called the clavam gene cluster due to the involvement of its members in 5*S* clavam metabolite biosynthesis. A pathway leading from clavaminic acid to the respective 5*S* clavam metabolites has been proposed by Baldwin *et al.* (1994), which involves oxidation, reduction, decarboxylation and hydrolysis activities (Figure 1.4.4). Therefore, some of the genes from the clavam gene cluster may encode enzymes that are involved in these steps, although the specific reactions catalyzed by each protein needs to be characterized, before any conclusions can be drawn. Jensen *et al.* (unpublished) have also prepared mutants defective in the *cvm*3, *cvm*6 and *cvm*7 and have shown that these mutations do not have any effect on either clavulanic acid or 5*S* clavam production. Therefore, their involvement in clavam metabolite biosynthesis remains questionable.

When the studies described in this thesis were initiated, only two clusters were known to comprise the genes involved in the biosynthesis of the clavam metabolites in *S. clavuligerus*. These were the previously described clavulanic acid (Jensen *et al.*, 1993; Jensen *et al.*, 2000) and clavam (Mosher *et al.*, 1999) gene clusters, respectively. Recently a second copy of the gene encoding PAH (Figure 1.4.1) was isolated and characterized in *S. clavuligerus* (Jensen *et al.*, 2004b). This second *pah* gene was not found to be located in either the clavulanic acid or the clavam gene clusters and was therefore called *pah*1 to differentiate it from *pah*2, which is located next to *cas*2 in the clavulanic acid gene cluster (Figures 1.4.2 and 1.4.3). In addition, the predicted amino acid sequences of PAH1 and PAH2 showed 71% identity or 81% similarity (Jensen *et al.*, 2004b). Jensen *et al.* (2004b) prepared *S. clavuligerus pah*1 single and *pah*1/*pah*2 double mutant strains and tested them for clavulanic acid and 5*S* clavam production. As predicted earlier by Jensen *et al.* (2000), the phenotypes of the *pah*1 and the *pah*1/*pa*h2 mutants were identical to the *cas*1 and the *cas*1/*cas*2 mutants (Table 1.2). This lent further support to the hypothesis that paralogues also exist for *cea*S, *bla* and *oat* as the phenotypes of the respective gene mutants were similar to those of the *cas*2 and *pah*2 mutants (Table 1.2). Results reported in this thesis and in related studies (Tahlan *et al.*, 2004c), have demonstrated the existence of the *cea*S, *bls*, and *oat* paralogues, which are grouped together with paralogues of other genes to form the paralogue gene cluster, which is the third gene cluster involved in the biosynthesis of the clavam compounds in *S. clavuligerus* (See Discussion for details).

Although the transcription of the genes involved in clavam metabolite biosynthesis has not been subjected to the same degree of rigorous investigation as that of the genes involved in cephamycin biosynthesis, studies suggest that most of the genes in the clavulanic acid gene cluster are also arranged in operons (Figure 1.4.2). The "early" genes from the clavulanic acid gene cluster that include *cea*S, *bls*, *pah*2 and *cas*2 are transcribed as a 5.3 kb polycistronic transcript, and *cas*2 is also transcribed as a 1.2 kb monocistronic transcript (Paradkar and Jensen, 1995) (Figure 1.4.2). In addition, Paradkar *et al.* (1995) observed that a transcript ~2.4 kb in size also hybridized to a *cas*2-specific probe during northern analysis of wild type *S. clavuligerus* RNA, and the identity of the additional ORF(s) present on this transcript is still not known. *oat* is transcribed as a 1.3 kb monocistronic transcript (Paradkar and Jensen, 1995) (Figure 1.4.2). Similarly, *opp*A1 is also transcribed as a monocistronic transcript (Paradkar *et al.*, 1998). During S1 nuclease protection studies carried out by Paradkar *et al.* (1998), transcripts spanning the *cas*2-*oat* and *oat*-*opp*A1 intergenic regions were also observed, indicating that additional uncharacterized polycistronic mRNA(s) might also encode *oat* and *opp*A1 (Figure 1.4.2). Both *cla*R and *cad* are transcribed as monocistronic messages (Perez-Redondo *et al.*, 1998), but Paradkar *et al.* (1998) have also suggested that a polycistronic transcript may contain both *cla*R and *cad* (Figure 1.4.2). A polycistronic transcript has been shown to encode *cyp* and *fd* (Li *et al.*, 2000)(Figure 1.4.2), providing further evidence that both of the encoded gene products may function together in a complex as suggested earlier.

The only other gene involved in clavulanic acid and 5*S* clavam metabolite biosynthesis subjected to transcriptional analysis is *cas*1, which is transcribed as a 1.4 kb monocistronic message (Paradkar and Jensen, 1995)(Figure 1.4.3). Therefore, a lot more information can be obtained by analyzing the transcription of other genes involved clavulanic acid and/or 5*S* clavam metabolite biosynthesis. In addition, the exact function of all the genes involved in clavulanic acid and 5*S* clavam metabolite biosynthesis needs to be elucidated, along with the identification of additional genes that may encode the missing enzymatic activities.

1.5 Regulation of β-lactam antibiotic biosynthesis in *S. clavuligerus*

Antibiotic production in streptomycetes is dependent on the growth phase and is linked to morphological development (Demain, 2000), both of which have many regulatory elements in common (Champness and Chater, 1994). In addition, a hierarchy of regulators coordinate both morphological development and secondary metabolism in streptomycetes (Horinouchi and Beppu, 1992a). At the top of this hierarchy are the global regulators that control both morphological development and secondary metabolism, whereas the pathway-specific transcriptional regulators that regulate the biosynthesis of one or more, closely related groups of secondary metabolites are thought to be the lowest members of this regulatory pyramid (Champness and Chater, 1994; Horinouchi and Beppu, 1992b). As well, growing evidence suggests that the pathways regulating antibiotic biosynthesis are not linear; instead, they form regulatory networks that may involve cross talk and the transmission of signals between different pathways (Horinouchi, 2003).

The *S̲treptomyces* A̲ntibiotic R̲egulatory P̲roteins (SARPs) contain N-terminal DNA binding domains similar to OmpR, and are pathway-specific regulators controlling the biosynthesis of one or more antibiotics (Wietzorrek and Bibb, 1997). The SARPs include the well-characterized ActII-ORF4 and RedD proteins that regulate the biosynthesis of the blue-pigmented polyketide actinorhodin (Act) and the red-pigmented tripyrolle undecylprodigiosin (Red) antibiotics in *S. coelicolor*, respectively (Fernandez-Moreno *et al.*, 1991; Narva and Feitelson, 1990). The third best-characterized member of the pathway-specific SARPs is probably DnrI from *S. peucetius*, which regulates the biosynthesis of the polyketide daunorubicin (Dnr) (Stutzman-Engwall *et al.*, 1992). The *cca*R (C̲ephamycin and C̲lavulanic A̲cid R̲egulator) gene from the cephamycin C gene cluster of *S. clavuligerus* also encodes a SARP, which not only regulates the biosynthesis of cephamycin C, but also that of clavulanic acid (Perez-Llarena *et al.*, 1997b). Therefore, CcaR can be considered to occupy a higher niche in the regulatory hierarchy in *S. clavuligerus*, as compared to ActII-ORF4, RedD and DnrI, which only regulate the biosynthesis of a single antibiotic in their respective organisms.

The SARPs are positive transcriptional regulators and are thought to bind to tandemly arrayed heptameric repeats, which are oriented on the same face of the DNA molecule, and may overlap with the promoter of the regulated gene(s) (Wietzorrek and Bibb, 1997). The consensus sequence of the predicted SARP binding repeats found in the Act and Dnr clusters is 5′-TCGAGCG/C-3′ and 5′-TCGAGCG-3′ (Wietzorrek and Bibb, 1997). In addition, DNase I footprinting has shown that both DnrI and ActII-ORF4 bind to sequences located within the -35 regions of the promoters they regulate, and the sequence of the binding site is 5'-TCGAG-3' (Arias *et al.*, 1999; Tang *et al.*, 1996). Since CcaR is also a SARP, it has been postulated that similar binding sites may also exist in *S. clavuligerus* for binding of CcaR (Santamarta *et al.*, 2002).

The gene encoding CcaR was first isolated and characterized by Perez-Llarena *et al.* (1997) and it was shown that a *cca*R mutant did not produce any detectable levels of either cephamycin C or clavulanic acid (Alexander and Jensen, 1998; Perez-Llarena *et al.*, 1997b). As well, the amplification of *cca*R in *S. clavuligerus* by the introduction of extra copies of the gene on a plasmid, led to a ~3 fold increase in cephamycin C and clavulanic acid production (Perez-Llarena *et al.*, 1997b). Northern analysis of wild type *S. clavuligerus* RNA revealed that the maximum levels of *cca*R transcription occur just prior to antibiotic production in *S. clavuligerus* (Bignell *et al.*, 2004; Perez-Llarena *et al.*, 1997b). In addition, western analysis revealed that maximum levels of the CcaR protein are also observed preceding antibiotic production (Bignell *et al.*, 2004). A similar trend was also observed in the case of *act*II-ORF4 and *red*D, where the expression of the respective genes preceded Act and Red production in *S. coelicolor* (Gramajo *et al.*, 1993; Takano *et al.*, 1992).

Over the years, many studies have been performed to examine the role of *cca*R in the regulation of genes involved in cephamycin C biosynthesis. By western analysis of cell free extracts (CFEs) from the *S. clavuligerus cca*R mutant, Alexander and Jensen, (1998) demonstrated that the CcaR, LAT, IPNS and DAOCS proteins were not expressed in the absence of CcaR (Figure 1.3.1). Due to the polycistronic nature of the *lat*, *pcb*AB and *pcb*C transcript (encoding LAT, ACVS and IPNS), it was proposed that CcaR either directly or indirectly controls the *lat* promoter (Figure 1.3.2). Therefore, by the same argument there should be no ACVS expression in the CcaR mutant also, as the *pcb*AB gene is part of the polycistronic *lat* transcript (Figure 1.3.2). In addition, since *pcb*C is also transcribed from its own independent promoter, the lack of IPNS in the *cca*R mutant indicated that CcaR somehow controls transcription from the *pcb*C promoter as well (Figure 1.3.2). The *cef*E gene, encoding DAOCS, is also transcribed as a polycistronic mRNA (Figure 1.3.2) suggesting that CcaR may regulate the upstream *cef*D promoter, which gives rise to the transcript in question (Figure 1.3.2). This is further evident from the fact that purified CcaR protein binds to the divergent *cef*D-*cmc*I promoters (Santamarta *et al.*, 2002) (Figure 1.3.2). Although Santamarta *et al.* (2002) showed that the activity of the *cef*D and the *cmc*I promoters increased ~1.7 fold in the presence of CcaR in a heterologous host, there is no additional information to suggest if CcaR is required for *cmc*I promoter activity in *S. clavuligerus*.

Even though the question has been addressed many times, it is still not clear if CcaR directly regulates the *lat* promoter. When CFEs from *E. coli* strains expressing His-tagged CcaR were used in Electrophoretic Mobility Shift Assays (EMSAs) along with a DNA fragment encompassing the *lat* promoter region, it was shown that some protein in the CFE could bind to the *lat* promoter (Kyung *et al.*, 2001). The same His-tagged CcaR-containing CFEs were used in DNase I footprinting analysis, and again it was shown that some protein could bind and protect the -35 region of the *lat* promoter, and the sequence of the deduced binding site was 5′-TCCAGC-3′. However, in another study when purified CcaR protein and the *lat* promoter region were used in EMSAs, the binding of CcaR to the *lat* promoter was not observed (Santamarta *et al.*, 2002). Therefore, it is still not clear at this stage, if CcaR directly binds upstream of *lat* to control expression, or if it exerts its effect indirectly on the *lat* promoter.

The CcaR protein binds to its own promoter and increases promoter activity ~4.4 fold when expressed in a heterologous host, implying positive self-regulation (Santamarta *et al.*, 2002). Overall, it is evident that CcaR positively regulates multiple operons involved in cephamycin C biosynthesis, either directly or indirectly (Figure 1.3.2). These operons include the "early" genes (*lat*→*pcb*C), the "intermediate" genes (*cef*D→*pcd*) and possibly the "late" genes also (*cmc*I→*cmc*J, Figure 1.3.2).

CcaR exerts its effect on clavulanic acid biosynthesis by controlling the expression of ClaR (Clavulanic Acid Regulator), which is the pathway-specific transcriptional regulator for clavulanic acid biosynthesis, and has been characterized by two independent groups (Paradkar *et al.*, 1998; Perez-Redondo *et al.*, 1998). ClaR belongs the LysR family of transcriptional regulators and contains two helix-turn-helix (HTH) DNA binding motifs at both its N- and C- terminal regions, and is therefore not a SARP. The *S. clavuligerus cla*R mutant does not produce any clavulanic acid, but still produces wild type levels of cephamycin C and the 5*S* clavams. In addition, overexpression of *cla*R in *S. clavuligerus* leads to a two-fold increase in clavulanic acid production (Perez-Redondo *et al.*, 1998). Therefore, it is thought that ClaR only regulates the late steps involved in the conversion of clavaminic acid to clavulanic acid, but not those involved in the biosynthesis of the 5*S* clavams (Figure 1.4.1). Paradkar *et al.* (1998) have shown that the expression of the *opp*A1, *cad* and *cyp* genes requires ClaR (Figure 1.4.2); therefore, these genes may represent "late" clavulanic acid biosynthetic genes. As well, ClaR does not control the expression of the "early" genes (*cea*S, *bls*, *pah*2, *cas*2 and *oat*, Figure 1.4.2), which are involved in the biosynthesis of both clavulanic acid and the 5*S* clavams (Figure 1.4.1). In addition, the transcription of *cla*R is reduced to near zero in a *S. clavuligerus cca*R mutant, resulting in the loss of clavulanic acid biosynthesis (Perez-Redondo *et al.*, 1998). In this manner, CcaR coordinates the biosynthesis of cephamycin C, a β-lactam antibiotic, with clavulanic acid, a β-lactamase inhibitor.

The pleiotropic antibiotic regulators come next in the regulatory hierarchy as they regulate the production of more than one secondary metabolite (Chater and Bibb, 1997), somewhat similar to the role of CcaR in *S. clavuligerus*. Although the exact functions of some of these regulators are not clear, they have been identified in more than one *Streptomyces* spp. These include *aba*A (Fernandez-Moreno *et al.*, 1992), *aba*B (Scheu *et al.*, 1997), *afs*Q_1/Q_2 encoding a sensor kinase/response regulator pair involved in signal transduction (Ishizuka *et al.*, 1992) and *afs*R (Horinouchi *et al.*, 1990). In *S. coelicolor*, AfsR has been shown to positively regulate secondary metabolism (Matsumoto *et al.*, 1994), whereas in *S. griseus* it is involved in regulating morphological differentiation, and not secondary metabolism (Umeyama *et al.*, 1999). The N-terminal region of AfsR is similar to the SARPs (Horinouchi *et al.*, 1990), and the production of all antibiotics is drastically reduced in a *S. coelicolor afs*R mutant (Floriano and Bibb, 1996). The region of the *S. coelicolor* chromosome surrounding *afs*R contains additional ORFs called *afs*S, *afs*K and *kbp*A (Umeyama *et al.*, 2002). The AfsK protein is similar to membrane associated eukaryotic serine/threonine kinases and has been shown to phosphorylate AfsR (Matsumoto *et al.*, 1994). *In vitro* studies have shown that KbpA can bind to the unphosphorylated form of AfsK and prevent autophosphorylation (Umeyama and Horinouchi, 2001). As well, overexpression of KbpA in *S. coelicolor* leads to a decrease in Act production, whereas a *kbp*A mutant overproduces Act, indicating that KbpA functions as a repressor, to indirectly decrease the levels of phosphorylated AfsR, consequently decreasing Act production (Umeyama and Horinouchi, 2001). PkaG and AfsL are two recently isolated serine/threonine kinases that are also able to phosphorylate AfsR, and a *S. coelicolor pka*G mutant was slightly compromised in its ability to produce Act (Sawai *et al.*, 2004). Therefore, it is possible that a number of serine/threonine kinases can phosphorylate AfsR in *S. coelicolor* (Petrickova and Petricek, 2003), to regulate the levels of Act produced in response to still unidentified external signals.

Overexpression of AfsS in *S. coelicolor* leads to the overproduction of Act and Red (Matsumoto *et al.*, 1995), and it has been shown that the phosphorylated form of AfsR binds to the *afs*S promoter, positively regulating its activity (Lee *et al.*, 2002). In addition, AfsS overexpression leads to an increase in the transcription of both *act*II-ORF4 and *red*D by a still unknown mechanism (Umeyama *et al.*, 2002). Therefore, the current hypothesis is that multiple kinases may phosphorylate AfsR, which in turn increases the transcription of *afs*S and additional genes that are involved in activating antibiotic biosynthesis (Horinouchi, 2003). It is not known if AfsR regulates either morphological differentiation or antibiotic biosynthesis in *S. clavuligerus*, although some type of regulatory role can be envisioned.

Some *Streptomyces* gene mutants are unable to form aerial mycelia and therefore form shiny/smooth colonies as compared to the wild type strain, which forms colonies that are fuzzy in appearance. Therefore, mutants unable to erect aerial mycelia are called bald (*bld*) mutants due to their appearance (Merrick, 1976). The *bld*A locus of *S. coelicolor* is probably the best characterized of all the *bld* genes, and *bld*A mutants are unable to form aerial mycelia or produce antibiotics, although vegetative growth is unaffected (Merrick, 1976). Therefore, *bld*A is a global regulator of both morphological development and antibiotic biosynthesis, and it encodes the only leucyl tRNA able to translate UUA codons, which are very rare in *Streptomyces* due to the high G+C content of their DNA (Lawlor *et al.*, 1987; Leskiw *et al.*, 1991b). The rare TTA codons are normally found in a few genes only, which are involved in aerial mycelium formation, antibiotic resistance or regulation, all of which are expressed later on in the growth cycle (Leskiw *et al.*, 1991a). Although the *bld*A gene is expressed throughout growth, it is post-transcriptionally regulated and the mature tRNA becomes more abundant as the culture ages (Leskiw *et al.*, 1993). In this manner, the translation of TTA-containing genes, which are required in the later stages of the life cycle, is regulated by *bld*A (Leskiw *et al.*, 1991a).

The *S. coelicolor bld*A mutant does not produce any Act as the gene for the pathway-specific regulator *act*II-ORF4 contains a single TTA codon. However, if the TTA codon is changed to the leucyl TTG codon, this leads to the production of Act in the *bld*A mutant (Fernandez-Moreno *et al.*, 1991). Although *red*D, which encodes that pathway-specific regulator for Red production, does not contain a TTA codon, the production of Red is still regulated by *bld*A (Narva and Feitelson, 1990). This is due to the presence of a *bld*A-dependent TTA codon in *red*Z, which encodes a response regulator that is essential for the expression of *red*D (White and Bibb, 1997).

In some instances, it has been shown that certain TTA codons can be translated in the absence of *bld*A in *S. coelicolor* (Leskiw *et al.*, 1991b). The *cca*R gene from *S. clavuligerus* also contains a single TTA codon (Alexander and Jensen, 1998). When a *S. clavuligerus bld*A mutant was prepared, it was found to be defective in aerial mycelium formation, but it still produced cephamycin C and clavulanic acid, and CcaR translation was also unaffected (Trepanier *et al.*, 2002). It was suggested that the TTA codon present in *cca*R was mistranslated in the *bld*A mutant, and that the mistranslation of the TTA codon was dependent on the nucleotide immediately adjacent to the 3′ end of the codon (Trepanier *et al.*, 2002). Based on compiled sequences of *bld*A-dependent and independent TTA codons, it has been proposed that mistranslated TTA codons are followed by G or A, whereas *bld*A-dependent codons are followed by C or T (Trepanier *et al.*, 2002). Therefore, since the TTA codon in *cca*R is immediately followed by a G, the biosynthesis of cephamycin C and clavulanic acid is not dependent on *bld*A in *S. clavuligerus* (Trepanier *et al.*, 2002).

Although the *bld*A mutation did not have any apparent effect on β-lactam production in *S. clavuligerus*, recently it was shown that mutagenizing the *bld*G locus in *S. clavuligerus* abolished antibiotic production and resulted in a bald phenotype (Bignell *et al.*, 2004). The BldG protein is thought to be an anti-anti-sigma factor that probably functions with an anti-sigma factor protein to regulate the activity of one or more sigma factors (Bignell *et al.*, 2000). Bignell *et al.* (2003) have demonstrated that a *S. coelicolor bld*G mutant is also unable to produce aerial mycelia or antibiotics, and that the activity of BldG is regulated by phosphorylation. In *S. clavuligerus*, it was shown that BldG somehow controls *cca*R expression as the transcription and translation of the *cca*R gene was completely abolished in the *bld*G mutant (Bignell *et al.*, 2004). Therefore, it is clear that BldG regulates β-lactam biosynthesis in *S. clavuligerus* via CcaR, but it is not known how BldG regulates CcaR expression.

Many *Streptomyces* spp. produce γ-butyrolactones (or butanolides), which are small diffusible signaling molecules that include A-factor from the streptomycin producer, *S. griseus* (Kholkov *et al.*, 1967). A-factor is produced in a growth-dependent manner until the end of the exponential phase, and at very low concentrations, it triggers both morphological differentiation and antibiotic production in *S. griseus* (Horinouchi and Beppu, 1994). The A-factor Receptor Protein (ArpA) is a transcriptional repressor and contains an N-terminal HTH DNA binding domain (Onaka *et al.*, 1995), which binds to conserved DNA motifs called autoregulatory elements (AREs)(Kinoshita *et al.*, 1999). One such ARE is present in the promoter of the *adp*A gene, whose expression is repressed by the bound ArpA protein. When A-factor is produced or is available, it binds to ArpA and causes the protein to dissociate from the *adp*A promoter, therefore permitting the transcription of *adp*A (Onishi *et al.*, 1999). The AdpA protein is a positive transcriptional regulator and is similar to the AraC/XylC family of proteins that contain a C-terminal HTH DNA binding domain (Onishi *et al.*, 1999). AdpA binds to the promoter of *str*R, which encodes the pathway-specific activator for streptomycin, and activates its transcription (Vujaklija *et al.*, 1993). As well, AdpA also binds to the promoters of additional genes involved in morphological development and activates their transcription

(Yamazaki *et al.*, 2004). Therefore, A-factor functions as a stimulus to signal the initiation of, and synchronize both antibiotic production and morphological differentiation in *S. griseus*.

A putative ARE element, which may be able to bind an ArpA-like protein, has also been identified upstream of the *cca*R gene in *S. clavuligerus* (Folcher *et al.*, 2001). In addition, a gene encoding a butyrolactone autoregulator receptor protein was cloned from *S. clavuligerus* recently (Kim *et al.*, 2004). The encoded ScaR (*S. clavuligerus* Autoregulator Receptor) protein was expressed in *E. coli* and was shown to bind to γ-butyrolactones. Therefore, it seems that autoinducers might control the expression of CcaR, which in turn controls cephamycin C and clavulanic acid biosynthesis in *S. clavuligerus*, although additional work is required to prove this hypothesis.

The genomes of *S. coelicolor* and *S. avermitilis* each show the presence of more than 60 ORFs encoding putative sigma factors (Bentley *et al.*, 2002; Ikeda *et al.*, 2003). The existence of multiple sigma factors in *S. coelicolor* has been known for some time now (Westpheling *et al.*, 1985), as four genes encoding σ^{70} type sigma factors (*hrd*A, *hrd*B, *hrd*C and *hrd*D) had already been identified (Tanaka *et al.*, 1988). Except for *hrd*B, the genes encoding σ^{hrdC} and σ^{hrdD} could be disrupted in *S. coelicolor*, indicating that σ^{hrdB} is likely an essential sigma factor (Buttner *et al.*, 1990). *In vitro* studies indicated that the transcription of *act*II-ORF4 and *red*D depended on the non-essential σ^{hrdD} (Fujii *et al.*, 1996). Later, when Buttner and Lewis (1992) prepared a triple mutant defective in *hrd*A, *hrd*C and *hrd*D, it still produced normal levels of Red and Act, indicating a redundancy in sigma factor function. In addition, *Streptomyces* promoter sequences display a very high degree of heterogeneity, suggesting the involvement of multiple sigma factors in gene expression (Strohl, 1992). There have been reported examples showing that a single gene may be transcribed from up to four promoters, which might be recognized by different RNA polymerase (RNAP) holoenzymes (Brown *et al.*, 1992; Buttner *et al.*, 1988). As well, in *S. antibioticus* it has been shown that the alternate sigma factor SigE controls the biosynthesis of actinomycin by a still unknown mechanism (Jones *et al.*, 1997), whereas in *S. coelicolor* a *sig*E mutant has no effect on antibiotic production (Paget *et al.*, 1999). Therefore, it is plausible that multiple and/or alternate sigma factors may be involved in regulating the biosynthesis of antibiotics in *S. clavuligerus* also.

The highly phosphorylated nucleotide (p)ppGpp is thought to play an important role in determining the onset of stationary phase and antibiotic production in *Streptomyces* (Ochi, 1986; Strauch *et al.*, 1991). The production of (p)ppGpp is associated with the "stringent response" resulting from amino acid starvation or nutrient downshifting, and (p)ppGpp functions as a global regulator to signal a change in growth phase (Cashel *et al.*, 1996). In *S. coelicolor* a positive correlation has been observed between the accumulation of ppGpp and the transcription of *act*II-ORF4 and *red*D (Strauch *et al.*, 1991; Takano *et al.*, 1992). When the ppGpp synthetase-encoding *rel*A gene was knocked out in *S. coelicolor*, antibiotic biosynthesis was abolished under nitrogen limiting conditions (Chakraburtty and Bibb, 1997). In addition, antibiotic production was completely abolished in a *S. antibioticus rel*A mutant (Hoyt and Jones, 1998). Recently two homologues of *rel*A (*rel*A and *rhs*) were also cloned and characterized in *S. clavuligerus* (Jin *et al.*, 2004). It was shown that the *S. clavuligerus rel*A mutant had a bald phenotype and it did not produce any detectable levels of β-lactam metabolites, whereas the *rhs* mutant was not so severely affected. Due to this reason, it was suggested that *rel*A was more important in terms of regulating morphological differentiation and antibiotic production in *S. clavuligerus*.

It is apparent that many interrelated factors are involved in regulating the biosynthesis of the β-lactam antibiotics in *S. clavuligerus*. As well, the regulatory pathways identified in other species suggest that many of them are conserved to some extent in most *Streptomyces*. External signals are received and transmitted via butyrolactone autoregulator receptor proteins, two component signal transducers and serine/threonine kinases, although the myriad of signals recognized by the latter two mechanisms are still unknown. In addition, homeostasis and internal signals also play a pivotal role in controlling antibiotic production. Both external and internal signals are then transmitted to the pathway specific regulators through additional mediators, to finally regulate and coordinate the biosynthesis of secondary metabolites. Therefore, a lot more work is required before some sort of rough network can be defined, which may be involved in regulating antibiotic biosynthesis in *Streptomyces*.

1.6 Main objectives of this thesis

The similarity in the phenotypes of the *cea*S, *pah*2 and *cas*2 mutants (Section 1.4 and Table 1.2) suggests that a second copy of *cea*S also exists in *S. clavuligerus*. Therefore, the main objective of this thesis was the isolation and characterization of the *cea*S paralogue. In addition, the region of the *S. clavuligerus* chromosome surrounding the putative *cea*S paralogue was analyzed to determine if paralogues of other genes, involved in the early stages of clavulanic acid and 5*S* clavam biosynthesis, are also clustered along with the putative *cea*S paralogue. Studies were carried out to determine if the clavulanic acid and the clavam gene clusters are physically linked. In addition, the location of the putative *cea*S paralogue relative to the clavulanic acid and the clavam gene clusters was also investigated. Lastly, the regulation of the paralogous pairs of genes involved in the early stages of clavulanic acid and 5*S* clavam metabolite biosynthesis was examined.

Of all the mutants defective in single copies of the genes involved in the early shared part of the pathway for which paralogues are postulated to exist, the *cea*S mutant seems to be the most severely compromised in its ability to produce clavulanic acid and the 5*S* clavams (Table 1.2), even though a second *cea*S gene probably exists in *S. clavuligerus* (Jensen *et al.*, 2000). It has been suggested that the phenotype of the *cea*S mutant may be due to polar effects on the expression of genes downstream of *cea*S, which are transcribed along with *cea*S on a polycistronic transcript (Jensen *et al.*, 2000). Therefore, the project described in this thesis was initiated with the preparation of new *cea*S mutants in an attempt to alleviate any associated polar effects, and major findings from the studies carried out are discussed in the following chapters.

CHAPTER 2: MATERIALS AND METHODS

2. Materials and Methods

2.1 Materials

Restriction and modification enzymes were obtained from New England Biolabs Ltd. (Mississauga, ON), Promega Corp. (Madison, WI) or Roche (Laval, QC). Dimethylsulfoxide (DMSO), Isopropyl-β-D-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-β-D-galactopyranoside (X-gal), N-tris-[hydroxymethyl]-methyl-2-aminoethane-sulfonic acid (TES) buffer, 3-[N-morpholino]-propanesulfonic acid (MOPS) buffer, imidazole, ampicillin (amp), hygromycin (hyg) kanamycin (kan), Penicillin G (PenG), spectinomycin (spec), tetracycline (tet) and thiostrepton (tsr) were purchased from Sigma Chemical Corp. (St. Louis, MO). Apramycin (apr) and chloramphenicol (cam) were purchased from PROVEL Division Eli Lilly Canada Inc. (Scarborough, ON) and Sigma-Aldrich Canada, Ltd. (Oakville, ON), respectively.

Polyethylene glycol (PEG) 1000 and 8000 were obtained from Koch-Light Ltd. (Haverhill, England). Nutrient agar (1.5%), Nutrient broth, Bacto™ peptone, Malt extract, ISP-4 medium, Bacto™ yeast extract, Bacto™ tryptone and soluble starch were all purchased from Difco Laboratories (Detroit, MI). DNA oligonucleotide primers used were synthesized by the Molecular Biology Service Unit, University of Alberta, or by Qiagen (Alameda, CA). All other chemicals used were reagent grade and were purchased from BDH Inc. (Toronto, ON) or EM Science (Darmstadt, Germany).

Microfuge tubes, Fisher brand® 0.5 ml for PCR and 1.5 ml G-Tube™ for RNA manipulation, were purchased from Fisher Scientific (Nepean, ON). Other (1.5 ml, 0.6 ml and 0.2 ml) microfuge tubes were obtained from Rose Scientific (Edmonton, AB). Quarter inch paper discs for bioassays were obtained from Schleicher and Schuell (Keene, NH). Fisher brand thick chromatography paper was purchased Fisher Scientific (Pittsburgh, PA). Kimax or Pyrex (USA) supplied most of the glassware used in this study.

All electrophoresis apparatus was purchased from Bio-Rad laboratories (Hercules, CA). Hybond™-N nylon membranes used for Southern, colony and northern hybridizations were purchased from Amersham Life Sciences Inc. (Oakville, Ontario).

Kodak X-OMAT AR film (Eastman Kodak Co., Rochester, New York) or a phosphorimager (Molecular Dynamics model 445 SI), was used for autoradiography. Western blots were exposed using Bioflex® MSI Film (Clonex Co.) and all X-ray films were developed using a Fuji RGII X-ray film processor (Fuji Photo Film Co., Tokyo).

A Branasonic 42-sonication bath (Branson cleaning equipment Co., Shelton, CN) was used for sonicating samples and an ATI UNICAM UV/Vis Spectrometer (UV3) was used to measure OD and absorbance values. The sources of all the other materials and enzymes are described along with their use in the subsequent sections.

2.2 Bacterial strains and plasmids

2.2.1 *Escherichia coli* and *Streptomyces* strains

Escherichia coli and *Streptomyces clavuligerus* strains used in this study are listed in Table 2.1 and Table 2.2, respectively. Other indicator organisms used in this study are described in the subsequent sections along with their specific use.

2.2.2 Plasmid vectors

Cloning vectors and recombinant plasmids used in this study are described in Table 2.3. A genomic library containing *S. clavuligerus* chromosomal DNA in the cosmid vector pWE15 (Table 2.3) was generously provided by Dr. W. Jin, Seoul National University, Seoul, Korea. Additional plasmids and cosmid clones used in linkage studies are described along with their use (Section 3.5.1).

2.3 Centrifugation and Harvesting of Cultures

Centrifugation of cultures was carried out using a Beckman Model J2-21 centrifuge. *E. coli* cultures were centrifuged at 5000 rpm using either a JA-20 or a JA-14 rotor, cooled to 4°C; whereas *S. clavuligerus* cultures were centrifuged at 8000 rpm using a JA-20 rotor. Unless otherwise mentioned, a clinical centrifuge (International Equipment, Needham, MS) at speed setting 7 was also used to harvest cultures and microfugation was performed using an Eppendorf 5415 C micro-centrifuge at 14000 rpm (Brinkmann Instruments Inc. Mississauga, Ontario).

Table 2.1: *E. coli* strains used in this study.

Strain	Description	Reference or Source
BW25113/pIJ790	Recombination host for REDIRECT© PCR targeting system, expresses recombination promoting proteins from plasmid pIJ790	Gust *et al*., 2003
DH5α	General cloning host	Gibco BRL
DH5α(BT340)	Host for the FLP-mediated excision of DNA flanked by FRT sites, expresses the FLP recombinase from the temperature sensitive plasmid BT340	Datsenko and Warner, 2000
ER1447	Methylation deficient strain used for isolating DNA for transformation into *Streptomyces*	J. McCormick, Harvard University, Cambridge, Mass. USA
ESS	β-lactam sensitive indicator strain	A. L. Demain, Massachusetts Institute of Technology, Cambridge. Mass. USA
ET12567	Methylation deficient strain used for conjugation and isolating DNA for transformation into *Streptomyces*	Kieser *et al*., 2000
HB101	Tetracycline sensitive host for pLAFR3 based clones	Salowe *et al*., 1990
One Shot™ TOP10	Host for TOPO™ cloning vectors	Invitrogen
XL1-Blue	General cloning host	Stratagene

Table 2.2: *S. clavuligerus* strains used in this study.

Strain	Description	Reference or Source
NRRL3585	Wild type strain producing cephamycin C, clavulanic acid and the 5*S* clavams	Northern Regional Research Laboratory, Peoria, Ill. USA
MEL	Mutagenized, easily transformable strain used for isolating plasmid DNA for transformation into other *S. clavuligerus* strains for preparing mutants	B. Barton, Glaxo SmithKline Pharmaceuticals, Worthing, UK.
4B and 4B-C	*cea*S2 insertional disruption mutants (*cea*S2::*apr*), apr^r	Jensen *et al.*, 2000
*cea*S2-Fs	*cea*S2 frame shift mutant	This study
*cea*S2-Fs-C	*cea*S2 frame shift mutant complemented with pSET152-5.3, apr^r	This study
*cea*S2-Fs-P	*cea*S2 frame shift mutant containing pSET152, apr^r	This study
Δ*cea*S2::*apr*	*cea*S2 deletion mutant prepared using the REDIRECT© technology, apr^r	This study
Δ*cea*S2	*cea*S2 inframe deletion mutant prepared using the FLP recombinase	This study
Δ*cea*S2-C	*cea*S2 inframe deletion mutant complemented with pSET152-5.3, apr^r	This study
Δ*cea*S2-P	*cea*S2 inframe deletion mutant containing pSET152, apr^r	This study

Table 2.2: (continued).

Strain	**Description**	**Reference or Source**
Δ*cea*S1::*apr*	*cea*S1 deletion mutant prepared using the REDIRECT$^{©}$ technology, apr^{r}	This study
Δ*cea*S1::*apr*/*cea*S2-Fs	*cea*S2 frame shift and *cea*S1 deletion mutant (*cea*S1/*cea*S2 double mutant), apr^{r}	This study
Δ*cca*R::*tsr*A	*cca*R deletion mutant; *cca*R replaced by *tsr*, tsr^{r}	Alexander and Jensen, 1998
CNG	Wild type *S. clavuligerus* contains a promotorless *egfp* reporter gene (on plasmid pIJ8660) integrated at the ϕC31 *attB* site in the chromosome, apr^{r}	This study
C1G	Wild type *S. clavuligerus* contains the *egfp* reporter gene fused to the *cea*S1 promoter; integrated at the ϕC31 *attB* site in the chromosome, apr^{r}	This study
C2G	Wild type *S. clavuligerus* contains the *egfp* reporter gene fused to the *cea*S2 promoter; integrated at the ϕC31 *attB* site in the chromosome, apr^{r}	This study
Δ*c7p*::*apr*	*c7p* deletion mutant prepared using the REDIRECT$^{©}$ technology; apr^{r}	This study
Δ*skn*::*apr*	*skn* deletion mutant prepared using the REDIRECT$^{©}$ technology, apr^{r}	This study

Table 2.3: Cloning vectors cosmids and recombinant plasmids used in this study.

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli* plasmids and cosmids**			
12B8-AP	Ampicillin, Apramycin, Kanamycin	Cosmid 12B8 having *cea*S2 deleted and replaced by *aac(3)IV* and the *oriT* of plasmid RP4	This study
pΔ12B8-AP (1, 2 and 3)	Ampicillin, Apramycin, Kanamycin	Cosmid 12B8-AP having different portions of the insert deleted but still containing the deleted *cea*S2 region encompassing *apr*	This study
6G9	Ampicillin, Kanamycin	pWE15-derived cosmid carrying a *S. clavuligerus* DNA fragment with *cea*S1, *bls*1, *pah*1 and *oat*1 and flanking sequences	Jensen et al., 2004b
8-51	Tetracycline	pLAFR3-derived cosmid clone carrying a *S. clavuligerus* DNA fragment encompassing most of *pcb*AB and downstream regions	Jensen *et al.*, (unpublished)
12B8	Ampicillin, Kanamycin	pWE15-derived cosmid carrying a DNA fragment encompassing *cea*S2 and the downstream region from the clavulanic acid gene cluster	This study
12B8-Δ*cea*S2	Ampicillin, Kanamycin	Cosmid 12B8 having *cea*S2 replaced by an in-frame deletion	This study
14E10	Ampicillin, Kanamycin	pWE15-derived cosmid carrying a *S. clavuligerus* DNA fragment with *cea*S1, *bls*1, *pah*1 and *oat*1 and flanking sequences	Jensen *et al.*, 2004b

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli* plasmids and cosmids**			
14E10-AP	Ampicillin, Apramycin, Kanamycin	Cosmid 14E10 having *cea*S1 deleted and replaced by *aac(3)IV* and the *oriT* of plasmid RP4	This study
14E10-C7P	Ampicillin, Apramycin, Kanamycin	Cosmid 14E10 having *c7p* deleted and replaced by *aac(3)IV* and the *oriT* of plasmid RP4	This study
14E10-SKN	Ampicillin, Apramycin, Kanamycin	Cosmid 14E10 having *skn* deleted and replaced by *aac(3)IV* and the *oriT* of plasmid RP4	This study
K6L2	Tetracycline	pLAFR3 based cosmid clone spanning the clavulanic acid gene cluster	Aidoo *et al.*, 1994
p2.8-18	Ampicillin	pUC18 containing a 2.8 kb *Eco*RI fragment carrying part of *cea*S1	This study
p2.8-18-Self	Ampicillin	p2.8-18 derivative where the region upstream of *cea*S1 was removed by *Sma*I/*Eco*47III digestion and religation	This study
p5N14E10B	Ampicillin	pUC120 containing a 5.2 kb *Nco*I fragment from cosmid 14E10 comprising *c7p* and *skn*	Jensen *et al.*, (unpublished)

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
E. coli **plasmids and cosmids**			
p5N14E10Sal	Ampicillin	pUC118 containing a 1.9 kb *Sal*I fragment from p5N14E10B	Jensen *et al.*, (unpublished)
p5.7	Ampicillin	pUC118 containing a 5.7 kb *Eco*RI fragment carrying the 3'-end of *cea*S1, all of *bls*1, *pah*1 and *oat1*, and flanking sequences	Tahlan *et al.*, 2004c
p200	Ampicillin	pUC120 containing the 200 bp *Nco*I fragment from p2.8-18	This study
p700	Kanamycin	pZErO$^{\text{TM}}$-2.1 containing a 700 bp *Eco*RI/*Apa*I fragment from p2.8-18	This study
p800-Sal	Ampicillin	pUC118 containing a 796 bp *Sal*I fragment from pO4H4. Carries the 3′-ends of *pah*1 and *oat*1 along with the intergenic region	Tahlan *et al.*, 2004b
pBB5.3A	Ampicillin	pUC119 containing the 5.3 kb *Bam*H1/*Bgl*II fragment encompassing *pcb*R and *cea*S2 from pTZ40	Tahlan *et al.*, 2004a
pBluescript II$^{®}$ KS^{+}	Ampicillin	High copy number *E. coli* phagemid cloning vector	Stratagene
pCAD2-3	Ampicillin	pUC120 containing a 4.3 kb *Eco*RI insert from K6L2, contains *cea*S2 and *bls*2	Jensen *et al.*, 2000

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli* plasmids and cosmids**			
pCAD2-3(L1-5)	Ampicillin	pCAD2-3 containing the frame shift copy of *cea*S2	This study
pCEC026	Ampicillin	pUC120 containing a 2.9 kb *Nco*I fragment carrying *cvm*1 and surrounding regions	Mosher *et al.*, 1999
pCR2.1®TOPO®	Ampicillin, Kanamycin	Phagemid for rapid cloning of PCR-generated fragments	Invitrogen
pDA115	Ampicillin	pBluescript II® KS$^+$ containing a 800 bp *Eco*RI/*Bam*HI fragment carrying parts of *rhs*A and *scl*U from the cephamycin gene cluster	Alexander and Jensen, 1998
pEB3	Ampicillin	pUC119 containing a 1.8 kb *Bam*HI/*Eco*RI DNA fragment encoding *orf*20 from the clavulanic acid gene cluster	Tahlan *et al.*, 2004b
pIJ773	Ampicillin, Apramycin	pBluescript KS+ derivative containing the apramycin resistance gene [*aac(3)IV*] and the *oriT* of plasmid RP4	Gust *et al.*, 2003
pJENII	Ampicillin	pUC119 containing a 3.6 kb *Kpn*I/*Sac*I DNA fragment encompassing the region downstream of *cvm*6 from the clavam gene cluster	Jensen *et al.*, (unpublished)

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli* plasmids and cosmids**			
pTOPO-*cea*S1-4	Ampicillin, Kanamycin	pCR2.1TOPO containing a 784 bp PCR fragment encompassing the *cea*S1 promoter	This study
pTOPO-*cea*S2-8	Ampicillin, Kanamycin	pCR2.1TOPO containing a 721 bp PCR fragment encompassing the *cea*S2 promoter	This study
pTZ40	Ampicillin	pTZ18R containing an approximately 26 kb *Bam*HI/*Pst*I fragment from the cosmid K6L2 spanning the clavulanic acid gene cluster	Tahlan *et al.*, 2004a
pUC18	Ampicillin	*E. coli* cloning vector	Stratagene
pUC18-0.9	Ampicillin	p2.8-18 having the 1.9 kb *Apa*I fragment deleted	This study
pUC118	Ampicillin	Phagemid cloning vector	Vieira and Messing, 1987
pUC120	Ampicillin	Phagemid cloning vector with a *Nco*I site in the MCS	Vieira and Messing, 1987
pUC120Apr	Ampicillin, Apramycin	pUC120 containing the *apr* gene flanked by *Nco*I sites	Paradkar and Jensen, 1995
pUZ8002	Kanamycin	Expresses RK2/RP4 transfer functions for conjugative transfer of plasmids containing the appropriate *oriT in trans*	Kieser *et al.*, 2000

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli* plasmids and cosmids**			
pWE15	Ampicillin, Kanamycin	*E. coli* cosmid cloning vector	Promega
pZErO^{TM}-2.1	Kanamycin	*E. coli* cloning vector	Invitrogen
pZErO-1.9	Kanamycin	pZErO^{TM}-2.1 containing the 1.9 kb *Apa*I fragment from p2.8-18	This study
***E. coli-Streptomyces* shuttle plasmids**			
pCAD2-3(L1-5)486	Ampicillin, Thiostrepton, Neomycin	*E. coli-S. clavuligerus* shuttle vector: pCAD2-3(L1-5) ligated to *Hind*III digested pIJ486	This study
pDA1006	Apramycin	pSET152 containing a 1.4 kb *Eco*RI/*Nru*I fragment encompassing *cca*R from the cephamycin gene cluster	Alexander and Jensen, 1998
pSET152	Apramycin	*E. coli* cloning vector, integrates into the *Streptomyces* chromosome	Northern Regional Research Laboratory, Peoria, Ill. USA
pSET152-5.3	Apramycin	pSET152 containing a 5.3 kb *Bam*HI/*Xba*I fragment from pBB5.3A encompassing *cea*S2 and upstream region	This study

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***E. coli-Streptomyces* shuttle plasmids**			
pUWL-KS	Ampicillin, Thiostrepton	High copy number *E. coli-Streptomyces* shuttle plasmid. Contains the pIJ101 replicon and the pBluescript KS MCS	Wehmeier, 1995
pUWL-Δ*ceaS2::apr*	Ampicillin, Apramycin, Thiostrepton	pUWL-KS containing the 5.5 kb *Kpn*I fragment from 12B8-AP encompassing Δ*cea*S2::*apr*	This study
pUWL-Δ*cea*S2	Ampicillin, Thiostrepton	pUWL-Δ*ceaS2::apr* having the disruption cassette replaced by an in-frame deletion in place of *cea*S2	This study
***Streptomyces* promoter probe plasmids**			
pIJ486	Thiostrepton, Neomycin	*Streptomyces* promoter probe plasmid used for subcloning DNA fragments in this study	Ward *et al.*, 1986
pIJ8660	Apramycin	Promoterless EGFP reporter plasmid, free living in *E. coli* and integrates into the *Streptomyces* chromosome	Sun *et al.*, 1999
pIJ8660-*cea*S2	Apramycin	EGFP reporter construct; pIJ8660 containing the 707 bp *Bam*HI/*Kpn*I fragment from pTOPO-*cea*S2-8 encompassing the *cea*S2 promoter region fused to the *egfp* gene	This study

Table 2.3: (continued).

Plasmid	Antibiotic Marker	Description	Reference or Source
***Streptomyces* promoter probe plasmids**			
pSET-*cea*S1	Apramycin	EGFP reporter construct; pSET152 containing the 2.37 kb *Bam*HI/*Eco*RI fragment from pTO6-*cea*S1 encompassing the *cea*S1 promoter fused to the *egfp* gene	This study
pT06	Spectinomycin	pIJ8860 derivative in which the apramycin resistance cassette (*aac(3)IV*) is replaced with the spectinomycin resistance gene (*aad*)	O'Connor *et al.*, 2002
pT06-*cea*S1	Spectinomycin	EGFP reporter construct; pT06 containing the 777 bp *Bam*HI/*Kpn*I fragment from pTOPO-*cea*S1-4	This study
pTO6-*cea*S2	Spectinomycin	EGFP reporter construct; pT06 containing the 707 bp *Bam*HI/*Kpn*I fragment from pTOPO-*cea*S2-8	This study

2.4 *E. coli* procedures

2.4.1 Culture conditions and media

E. coli strains were grown at 37°C unless otherwise indicated. Liquid cultures were grown in either Lennox broth (LB) (1% tryptone, 0.5% yeast extract, 0.5% NaCl and 0.1% glucose) or SOB medium (Sambrook *et al.*, 1989) using a Cell Production Roller Drum (Bellco Biotechnology) or a G-24 Environmental incubator shaker (New Brunswick Scientific CO.). Solid LB medium containing 1.5% agar was used to maintain plate cultures. Plasmid-bearing *E. coli* cultures were supplemented with ampicillin (100 μg/ml), apramycin (50 μg/ml), kanamycin (30 μg/ml), chloramphenicol (25 μg/ml), tetracycline (10-12.5 μg/ml) or spectinomycin (100 μg/ml).

2.4.2 Preparation of *E. coli* glycerol stocks

E. coli cultures were stored as glycerol stocks. One milliliter amounts of overnight *E. coli* cultures grown in LB were microfuged and the supernatant was removed by aspiration. The cell pellets were resuspended in 500 μl of sterile 20% glycerol and the stocks were then stored at –80°C.

2.4.3 Preparation of electrocompetent *E. coli* cells

Ten microliter amounts of the appropriate *E. coli* glycerol stocks were used to inoculate 10 ml of LB, containing the suitable antibiotic(s), and the tubes were incubated overnight at 37°C. Overnight cultures (200 μl) were used to inoculate 200 ml of SOB containing 200 mM $MgSO_4$ in a 500 ml flask, and the cultures were incubated with shaking at 37°C for 3-4 hours or until the OD_{600} was ~ 0.4-0.6. The cultures were split up into two 100 ml aliquots, which were chilled on ice and harvested at 4°C. The cells were washed twice in 100 ml of ice-cold 10% glycerol, which was prepared using Milli-Q™ water. Care was taken to keep the cells and solutions chilled at all times, and any surface coming in contact with the cells was washed multiple times with Milli-Q™ water before sterilization. After the washes, each 100 ml aliquot of the original culture was resuspended in 4 ml of 10% glycerol, and 40 μl aliquots of the cell suspension were dispensed into 1.5 ml microcentrifuge tubes. The cells were then used for electroporation immediately or were flash frozen in a dry ice ethanol bath, and were stored at –80°C for future use.

E. coli DH5α/BT340 electrocompetent cells were prepared using the same procedure as mentioned above except that all incubations were carried out at 30°C.

E. coli BW25113/pIJ790 electrocompetent cells containing the Cosmids 14E10 (used in the PCR-targeted mutagenesis of *cea*S1, *c7p* and *skn*) and 12B8 (used in the PCR-targeted mutagenesis of *cea*S2) were also prepared using the same procedure mentioned above, with some changes. In addition to antibiotics (ampicillin, kanamycin and chloramphenicol), the culture media also contained 10 mM L-arabinose, which induces the expression of the λRED proteins. The main cultures were 100 ml in volume, and after multiple washes with 10% glycerol, the cells were resuspended in a final volume of 200 μl of 10% glycerol, 50 μl of which was then immediately used for transformation.

2.4.4 Electroporation of DNA into electrocompetent *E. coli*

Ligation reactions were precipitated and the DNA pellets were redissolved in 10 μl of sterile Milli-Q™ water, after washing with 70% ethanol to remove as much salt from the samples as possible. Five microliter amounts of the redissolved ligation reactions were used to transform 40 μl of electrocompetent *E. coli* cells. Previously prepared frozen electrocompetent cells were slowly thawed on ice before use. Two millimeter gap BTX® Disposable Cuvettes Plus™ (Genetronics, Inc., San Diego, CA) were used for electroporation along with a Bio-Rad GenePulser II apparatus set to 25 μF capacitance, 200 Ω resistance and 2.5 kV, and both the cuvettes and the holding apparatus were cooled to 4°C before use. Immediately after the pulse, 955 μl of ice cold LB was added to the cells and 1 ml of the cell suspension was gently transferred to a 1.5 ml microfuge tube. The resuspended cells were then allowed to recover at either 37 or 30°C for 1 to 2 hours, depending on the plasmid and electrocompetent *E. coli* strain used in the transformation. Aliquots of 100 μl and 900 μl (after reduction to 100 μl) were then plated onto LB agar containing the appropriate antibiotic(s) at room temperature. In some cases, IPTG (100mM) and X-gal (40 μg/ml) were also included in the medium to screen for the presence of inserted DNA in specific vectors, based on α-complementation by blue-white selection. The plates were incubated overnight at the appropriate temperature for the isolation of transformants.

2.4.5 Isolation of plasmid and cosmid DNA from *E. coli*

Plasmid and cosmid DNA was isolated from *E. coli* cultures using the Birnboim and Doly method (1979) described by Sambrook *et al.* (1989).

2.4.6 Preparation of colony blots

Using sterile toothpicks, plasmid or cosmid-containing *E. coli* colonies were transferred onto circular Hybond™-N membranes placed on top of LB agar containing the appropriate antibiotics. The plates were incubated overnight at 37°C, after which the membranes were carefully removed and treated to lyse the cells and release plasmid/cosmid DNA, as described by Sambrook *et al.* (1989). Briefly, the membranes were placed, colony side up for 5 minutes, onto filter papers soaked in (1) 10% SDS, (2) denaturation solution (0.5 M NaOH and 1.5 M NaCl), (3) neutralization solution [1.5 M NaCl and 0.5 M Tris-HCl (pH 7.4)] and (4) 2× SSC (0.3 M NaCl and 0.03 M tri-sodium citrate). DNA was fixed onto the membranes by UV exposure (150 mJoules) followed by baking for two hours in a vacuum oven at 80°C and 7.3 psi. The membranes were then treated with Proteinase K (Roche) at a final concentration of 50 µg/ml, overnight at 50°C with gentle agitation in a buffer consisting of 0.01 M Tris (pH 7.8), 0.005M EDTA and 0.5% SDS. They were then washed twice in 2× SSC to remove any residual cell debris before being stored at room temperature.

2.5 *Streptomyces* procedures

2.5.1 Culture conditions and media

S. clavuligerus strains were routinely grown in 25 ml of liquid culture media in 125 ml flasks. Strains were maintained on either MYM (0.4% maltose, 0.4% yeast extract, 1.0% malt extract and 1.8% agar) (Stuttard, 1982.) or ISP-4 (Difco) plates. Additional media used in this study are listed in the subsequent sections as per their use. Plasmid-bearing *Streptomyces* cultures were supplemented with apramycin (25 µg/ml), kanamycin (50 μg/ml) or thiostrepton (5 μg/ml for *S. clavuligerus* and 50 μg/ml for *S. lividans*). All *Streptomyces* cultures were grown at 28°C and liquid cultures were grown on a rotary shaker (Model G-25, New Brunswick Scientific Co.) at 250 rpm.

2.5.2 Storage of strains

Streptomyces spore and mycelial stocks were prepared by growing strains on ISP-4 medium for ~7 days. Spore stocks were prepared by scraping the growth from one plate into 1 ml of 20% glycerol in a microfuge tube, which was then placed in a sonication bath for 1 minute to disperse the spores. The spore suspension was filtered through non-absorbent sterile cotton wool placed in a 10 ml syringe, and the spores were collected in a microfuge tube. Mycelial stocks were prepared using the same procedure described above, except that the sonication and the filtration steps were omitted. The spore and mycelial stocks were then stored at –80°C in microfuge tubes.

2.5.3 Preparation of *S. clavuligerus* protoplasts

S. clavuligerus spore stocks were used to inoculate 25 ml of culture medium consisting of 15 ml modified YEME [(0.3% yeast extract, 0.5% tryptone-peptone, 0.3% malt extract, 34% sucrose and 1% maltose) + 5 mM $MgCl_2.6H_2O$] and 10 ml of TSBM [tripticase soy broth (TSB) + 1% maltose] supplemented with 0.5% glycine and the appropriate antibiotic(s). The cultures were propagated for 30-40 hours after which they were harvested by centrifugation and washed twice in 10.3% sucrose. The mycelia were either frozen at –20°C for future use, or were used for preparing protoplasts immediately. If frozen mycelia were used for preparing protoplasts, the samples were washed once in 10.3% sucrose before use. The washed mycelia were resuspended in 4 ml of P buffer [10.3% sucrose, 0.057 mM K_2SO_4, 10 mM $MgCl_2$, 0.005% KH_2PO_4, 0.368% $CaCl_2$, 0.573% TES (pH 7.2), 0.59 μM $ZnCl_2$, 1.48 μM $FeCl_3$, 0.12 μM $CuCl_2$, 0.10 μM $MnCl_2$, 0.052 μM $Na_2B_4O_7$ and 0.016 μM $(NH_4)_6Mo_7O_{24}$] (Kieser *et al.*, 2000) + 2 mg/ml lysozyme and the samples were incubated at 30°C for 30 minutes. The tubes were gently rotated every 5-10 minutes to resuspend the mycelia. Progression of the protoplasting of the mycelia was followed every 10 minutes using a phase contrast microscope. Once protoplasting was complete, 5 ml of additional P buffer was added to the samples and the protoplasts were filtered through a sterile non-adsorbent cotton wool plug placed in a Pasteur pipette. The protoplasts were harvested and washed twice in 10 ml of P buffer using a clinical centrifuge (speed setting 6, ~3000 rpm; International Equipment, Needham, MS), and were resuspended in 5 ml of P buffer. Aliquots (1 ml) were then dispensed into 1.5 ml microfuge tubes and were used immediately for transformation, or were slowly frozen for future use by placing the tubes in a container of ice, which was then placed at –70°C.

2.5.4 Transformation of *S. clavuligerus* protoplasts

Transformation of *S. clavuligerus* protoplasts was carried out using previously described procedures with some modifications (Bailey and Winstanley, 1985; Kieser *et al.*, 2000). All centrifugation steps involving protoplasts were carried out at 3000 rpm for 7 minutes at room temperature, using a microcentrifuge. Tubes containing frozen protoplasts were thawed quickly and the protoplasts were recovered by microfugation. The protoplasts were washed once in 1 ml of P buffer, which was removed gently using a pipette. The protoplasts were resuspended in the residual buffer by flicking the tube very gently, and were then heat shocked at 45°C for 5 minutes by placing the tubes in a water bath. Plasmid DNA in a maximum volume of 20 µl in either TE buffer [10 mM Tris-HCl (pH 8.0) and 1 mM EDTA (pH 8.0)] or Milli-Q™ dH_2O was then added to the protoplasts, followed immediately by the addition of 100 µl of T buffer (P buffer + 25% PEG 1000). After 1 minute at room temperature, 1 ml of P buffer was added to the tubes, which were then microfuged to recover the protoplasts. The pellets were gently resuspended in 500 µl of P buffer and 100 µl aliquots were plated out onto plates containing 20 ml of modified R5B medium [10% sucrose, 1% soluble salt, 0.1% casamino acids, 65 mM sodium glutamate, 0.2 mM $MgSO_4$, 25 mM $MgCl_2$, 0.005% KH_2PO_4, 0.368% $CaCl_2$, 0.573% TES (pH 7.2), 0.59 µM $ZnCl_2$, 1.48 µM $FeCl_3$, 0.12 µM $CuCl_2$, 0.10 µM $MnCl_2$, 0.052 µM $Na_2B_4O_7$, 0.016 µM $(NH_4)_6Mo_7O_{24}$ and 2.7% agar] (Bailey and Winstanley, 1985), which had been dried for ~2 hours.

DNA used for transforming *S. clavuligerus* protoplasts was isolated from methylation-deficient strains of *E. coli* such as ER1447 or ET12567, or from *S. clavuligerus* MEL, to prevent restriction of foreign DNA in the transformed protoplasts.

In certain cases, protoplasts were transformed with alkali-denatured plasmid DNA as described previously (OH and Chater, 1997). Briefly, 10 ml cultures of *E. coli* were used isolate plasmid DNA, which was redissolved in 9 µl of Milli-Q™ H_2O. A 2 µl amount of 1 M NaOH was added to the DNA solution, which was then incubated at 37°C for 10 minutes. Then 2 µl of 1 M HCl was added to the tube to neutralize the alkali, and the entire denatured plasmid preparation was used for protoplast transformation.

Transformed protoplasts were allowed to regenerate for 36-48 hours before the plates were overlaid with 3 ml of soft nutrient agar (0.8% nutrient broth and 0.7% agar) (Kieser *et al.*, 2000) containing the appropriate antibiotic, which was present in the nutrient agar at a concentration suitable for 23 ml final volume (volume of plate and overlay). Once the nutrient agar solidified at room temperature, the plates were returned to the incubator and after 3-5 days, transformants were normally observed on the plates.

Colonies arising from primary transformants on modified R5B plates were patched onto MYM agar plates containing the appropriate antibiotic(s). After 3-6 days of incubation, MYM plates with growth were assumed to represent true transformants, and growth from these plates was used to patch ISP-4 agar plates for the preparation of spore stocks. In some cases when good growth on the ISP-4 agar plates was not observed, mycelia from the MYM plates were scraped into microfuge tubes containing 100 µl of 20% glycerol. The tubes were sonicated in a sonication bath for 30 seconds and the entire contents were plated out onto ISP-4 plates. This procedure aided in the breaking up and spreading of mycelia, yielding better growth, and was used on numerous occasions to improve the quality of the inoculum used for preparing stocks.

2.5.5 Introduction of DNA into *S. clavuligerus* by inter-generic conjugation

Plasmid DNA was introduced into *S. clavuligerus* as described by Kieser *et al.* (2000) except that AS-1 medium [0.1% yeast extract, 0.02% L-alanine, 0.02% L-arginine, 0.05% L-asparagine, 0.5% soluble starch, 0.25% NaCl, 1% Na_2SO_4, 2% agar (pH 7.5)] (Baltz, 1980) supplemented with 10 mM $MgCl_2$ was used for isolating exconjugants. Briefly, *E. coli* ET12567 containing the plasmid pUZ8002 as well as the second plasmid to be mobilized was cultured overnight in LB with the appropriate antibiotics. The overnight culture (1% or 100 µl) was used to inoculate 10 ml of fresh, antibiotic-containing LB medium, and the *E. coli* culture was incubated at 37°C until the OD_{600} of the culture reached ~ 0.4-0.6. The cells were then harvested and washed twice in 10 ml of LB to remove most of the antibiotics present in the culture, and were resuspended in 600 µl of LB medium. Meanwhile, 100 µl of the appropriate *S. clavuligerus* spore stock was mixed with 500 µl of 2YT broth (Sambrook *et al.*, 1989) and was heat shocked at 50°C for 10 minutes. After cooling to room temperature, the heat-shocked spores were mixed with 500 µl of the *E. coli* cell suspension. The conjugation mix was microfuged to pellet the spore and cell mixture, 900 µl of the supernatant was removed, and the cell pellet was resuspended in the remaining 200 µl of culture media. One hundred microliter aliquots of the conjugation mix were plated onto 25 ml of AS-1 medium plates, which were then incubated at 28°C. After 16-20 hours of incubation, the plates were overlaid with 1 ml of sterile water containing 0.5 g nalidixic acid to give a final concentration of 20 µg/ml, and either apramycin or hygromycin to give final concentrations corresponding to 25 and 150 µg/ml, respectively, depending on the antibiotic marker used. The plates were further incubated for up to 8 days at 28°C, by which time colonies were normally observed.

2.5.6 Plasmid DNA isolation

Plasmid DNA from *S. clavuligerus* was isolated using a modified version of the Birnboim and Doly method (Birnboim and Doly, 1979), designed specifically for *Streptomyces*. Plasmid-bearing cultures were grown in TSB + 1% starch (TSBS), 1.5 ml of each culture was harvested by microfugation and the supernatant was removed by aspiration. Two hundred microliters of solution A [50 mM glucose, 25 mM Tris (pH 8.0), 10 mM EDTA and 2 mg/ml lysozyme] was added to the tubes and the mycelial pellets were resuspended by vortexing. The tubes were incubated at 37°C for 30 minutes and then 400 µl of freshly prepared solution B (0.2 N NaOH, 1% SDS) was added to the tubes and the contents were mixed by inverting the tubes multiple times. The tubes were placed on ice for at least 5 minutes, after which 300 µl of solution C [3 M potassium acetate (pH 4.8)] was added and the contents were mixed again as described earlier. The tubes were incubated on ice for at least 5 minutes and were then microfuged for 4 minutes. Eight hundred microliters of each supernatant was transferred to a fresh tube and was extracted with 200 µl of buffered phenol/chloroform/isoamyl alcohol (25:24:1) (Kieser *et al*., 2000). After microfugation, 750 µl of the aqueous phase was transferred to a fresh tube containing 750 µl of isopropanol and the tubes were placed on ice for 5-10 minutes. The DNA was pelleted by microfugation for 10 minutes and was redissolved in 50 µl of 100 mM ammonium acetate. One hundred and fifty microliters of 98% ethanol (−20°C) was then added to the samples, and the tubes were vortexed and microfuged as before. The recovered pellets were washed in 250 µl of 70% ethanol, air dried and redissolved in 25 µl of TE buffer with or without 50 µg/ml RNaseA.

2.5.7 Isolation of gene replacement mutants

Two strategies were used for isolating *S. clavuligerus* mutants prepared by homologous recombination. In the first case, mutated genes carried on *E. coli-Streptomyces* shuttle plasmids based on pIJ101 derivatives, which are easily lost from *Streptomyces* (Ward *et al*., 1986; Wehmeier, 1995), were transformed into *S. clavuligerus* protoplasts. In the second strategy, mutated genes were carried on cosmids, which could be mobilized into *S. clavuligerus* by conjugation, but were unable to replicate, and therefore functioned as suicide vectors (Gust *et al*., 2003).

Transformants obtained by protoplast transformation were tested, based on antibiotic resistance on MYM medium plates, to verify if they contained the plasmid of interest. These primary transformants were then patched onto ISP-4 plates without any antibiotic selection, to promote homologous recombination and the loss of the plasmid vector. The desired mutants were isolated using a previously described protocol (Aidoo *et al*., 1993). After the first round of sporulation on nonselective ISP-4 medium, spore stocks were prepared as described in section 2.5.2. Dilution series were set up using sterile 0.1% peptone or water and were plated out onto nonselective ISP-4 medium to give approximately 100-200 colonies per plate (this usually corresponded to dilutions ranging from 10^{-6}-10^{-7}). After ~7 days of growth, when sporulating colonies were observed on the ISP-4 plates, replica plating was carried out onto MYM medium. The order of platings onto antibiotic-containing medium depended on the predicted phenotype of the desired mutant. Normally the first replica plating was onto MYM master plates without any antibiotic. The second and third replica platings were carried out onto plates containing antibiotics specific for the disruption marker and the marker present on the plasmid vector, respectively. Colonies that were resistant to the antibiotic used as the disruption marker and sensitive to the one present on the plasmid vector, were presumed to have undergone double homologous recombination with the loss of the plasmid vector, and were subjected to further analysis. In some cases when frame-shifted or in-frame deletion mutants were prepared, the parent strain was already disrupted in the respective gene using an antibiotic resistance gene cassette. In these cases, the loss of the antibiotic resistance gene cassette disrupting the gene of interest was followed to isolate the desired mutant.

A different protocol was used to isolate mutants when suicide cosmids were used to introduce foreign DNA into *S. clavuligerus* by conjugation. Exconjugants were selected on AS-1 medium based on the resistance cassette (apr^r) used for gene disruption. The exconjugants that had undergone double cross-over events were selected by patching them onto MYM plates containing the antibiotic marker used in gene disruption (*apr*) and the antibiotic marker present on the cosmid vector (*kan*). Mycelia from exconjugants that were apr^r and kan^s were patched onto ISP-4 plates without selection and spore stocks were prepared as described earlier. Dilution series were set up to isolate unigenomic spores that gave rise to apr^r colonies, which were subjected to further analysis.

2.5.8 Fermentation of *S. clavuligerus*

Strains of *S. clavuligerus* were grown in fermentation medium to determine the levels of cephamycin C, clavulanic acid and 5*S* clavam production. Spore stocks (10-20 µl) were used to inoculate TSBS seed cultures, which were propagated for 30-48 hours or until the OD_{600} reached ~10-15. The TSBS seed cultures were then used to inoculate different fermentation media to 2% v/v. The compositions (per liter) of the fermentation media used in this study were as follows: starch asparagine (SA) [10 g soluble starch, 2 g asparagine, 21 g MOPS, 0.6 g $MgSO_4.7H_2O$, 4.4 g K_2HPO_4 and 1 ml trace elements solution (0.1% $FeSO_4.7H_2O$, 0.1% $MnCl_2.4H_2O$, 0.1% $ZnSO_4.7H_2O$ and 0.13% $CaCl_2.3H_2O$), adjusted to pH 6.8 with NaOH] (Aharonowitz and Demain, 1978); soy (15 g soybean flour, 47 g soluble starch, 0.1 g KH_2PO_4 and 0.2 g $FeSO_4.7H_2O$, adjusted to pH 6.8 with NaOH) (Salowe *et al.*, 1990) and glycerol sucrose proline glutamic acid (GSPG) (15 g glycerol, 20 g sucrose, 2.5 g proline, 1.5 g glutamic acid 5 g NaCl, 2 g K_2HPO_4, 0.4 g $CaCl_2$, 0.1 g $MnCl_2.4H_2O$, 0.1 g $FeCl_3.6H_2O$, 0.05 g $ZnCl_2$ and 1 g $MgSO_4.7H_2O$, adjusted to pH 7) (Romero *et al.*, 1984). In some fermentation experiments, the media were supplemented with glycerol (15 mg/ml) or dehydroacetic acid (DHA) (0.5 mg/ml). Unless otherwise specified, the cultures were harvested after 72 and 96 hours of growth by centrifugation. The culture supernatants were used for analysis of antibiotic production as described in sections 2.6.2, 2.6.3, 2.6.4 and 2.6.5, whereas the cell pellets were used for determining growth levels as described in section 2.6.1.

2.5.9 Isolation of chromosomal DNA

A modified version of the previously described procedure for isolating genomic DNA from *Streptomyces* was used in this study (Hopwood *et al.*, 1985). Fifteen-milliliter amounts of *S. clavuligerus* cultures grown in TSBS were harvested by centrifugation and the mycelia were washed twice in 10.3% sucrose. The mycelia were used immediately for genomic DNA preparation or were frozen at –20°C for future use. One milliliter of the harvested mycelial pellet was transferred to a 13 ml polypropylene tube and was resuspended in 4 ml of lysozyme buffer [25 mM Tris-HCl (pH 8.0), 0.3 M sucrose, 25 mM EDTA, 4 mg/ml lysozyme and 50 µg/ml RNaseA]. The tube was incubated at 37°C for 30-45 minutes, after which 1 ml of 2% SDS was added followed by vigorous shaking, until the viscosity of the contents decreased. One milliliter of phenol/chloroform/isoamyl alcohol was added to the tube, which was then vortexed vigorously for 1 minute. The tube was centrifuged at 12100 × *g* for 5 minutes and the aqueous phase was transferred to a fresh tube, and was re-extracted with phenol/chloroform/isoamyl alcohol one more time. Then the aqueous phase was extracted twice with an equal volume of chloroform and after the last extraction, 0.1 volumes of 3 M sodium acetate (pH 5.2) and either 1 volume of isopropanol or 2 volumes of 98% ethanol were layered on top of the aqueous phase. The tube was inverted gently to mix the two phases and was then placed on ice for 10 minutes. The tube was vortexed, and the clump of chromosomal DNA formed was transferred to a microfuge tube containing 70% ethanol. The DNA was pelleted for 10 minutes, air-dried and was allowed to redissolve overnight at 4°C in 100 µl of TE buffer containing 50 µg/ml RNaseA. In some cases, 400 µl of TE buffer was added to the sample the next day followed by one extraction with an equal volume of phenol/chloroform/isoamyl alcohol and one extraction with an equal volume of chloroform. The chromosomal DNA was then ethanol precipitated and redissolved in TE buffer.

2.5.10 Preparation of agarose plugs containing chromosomal DNA

Agarose plugs containing *S. clavuligerus* chromosomal DNA for PFGE were prepared as described earlier (Evans and Dyson, 1993), with some minor changes. Cultures of *S. clavuligerus* were grown in TSBS for 30-36 hours. Ten milliliters of each culture were harvested and washed twice in 10.3% sucrose and the mycelia were then resuspended in 1.5 mL of 4.0 M guanidine hydrochloride + 0.2 M EDTA in HE buffer [10mM HEPES-NaOH (pH 8.0), 1mM EDTA], and incubated at 37°C for 5 minutes. The mycelia were harvested again and resuspended in 1 mL of HES buffer [25 mM HEPES-NaOH (pH 8.0), 25 mM EDTA and 0.3 M sucrose]. The resuspended mycelia were mixed with an equal volume of molten 1.5% low melting agarose at 50°C and formed into plugs. The agarose plugs were allowed to set at 4°C for 15 minutes after which they were treated with 10 ml of HES buffer containing 10-20 mg/ml lysozyme at 37°C for 2-3 hours. The plugs were then treated with Proteinase K (Roche) overnight at 50°C with gentle agitation, in a tube containing 5 ml NDS buffer [1% SDS, 0.5 M EDTA and 10 mM glycine (pH 9.5)]. The following day, the plugs were treated with 3 mL of HE buffer containing 10 mg/ml Pefabloc® SC (AEBSF) protease inhibitor (Roche) at room temperature for 1 hour. They were then washed 4 times in 20 ml of HE buffer with gentle shaking. Each wash was 30 minutes at room temperature except for the second wash, which was carried at 37°C. The plugs were then stored at 4°C in HE buffer for up to 3 months.

2.5.11 Isolation of RNA

Two methods were utilized to start *S. clavuligerus* cultures for isolating RNA. For RNA used in S1 nuclease and primer extension analyses, 15 µl of a wild type *S. clavuligerus* spore stock containing ~10^9 spores/ml was used to inoculate 25 ml TSBS cultures, which were allowed to grow for 30-36 hours. Five hundred microliters of the TSBS cultures were used as the inocula to start 25 ml soy cultures, which were harvested after 96 and 120 hours of growth and were processed for RNA isolation as described below. For RNA used in RT-PCR and northern analysis, three-milliliter amounts of spore stocks (containing ~10^9 spores/ml) were heat shocked at 50°C for 10 minutes and were then pre-germinated for 4 hours at 28°C with shaking, in 15 ml of 2YT broth containing 0.5% glycerol. The spores were then harvested, and 1 ml and 0.5 ml amounts of the pre-germinated spores were used to inoculate 100 ml of SA and soy culture medium, respectively. RNA was isolated after 96 and 120 hours of growth in soy, and 72 and 96 hours of growth in SA medium.

The RNA used in this study was isolated using the Modified Kirby procedure (Kirby *et al.*, 1967) as described by Kieser *et al.* (2000). The cultures were harvested by filtration through a Whatman #1 filter paper (Whatman International Ltd., England) and the mycelia were scraped directly into chilled universal bottles (United Glass, United Kingdom), containing approximately 14 g of 3 mm sterile glass beads (Fisher) and 5 ml 1× Kirby's mix [1% sodium tri-isopropylnaphthalene sulfonate, 6% 4-amino salicylate and 6% neutral phenol buffered in 50 mM Tris-HCl (pH 8.3)]. Normally, SA and early time point soy cultures were divided into two samples at this stage, whereas the later soy cultures were divided into four samples, and in both cases, the samples for a particular time point were re-combined after DNAse treatment (see below). The bottles were vortexed vigorously 4-5 times (30-45 seconds each time), making sure that they were kept cool by intermittingly placing them on ice. Five milliliters of phenol/chloroform was then added and the samples were vortexed for an additional 3-4 minutes as mentioned above. At this stage the samples were frozen at –20°C, until all the cultures were harvested and processed together. The samples were then carefully transferred to 13 ml polypropylene tubes, which were then centrifuged at 8740 × *g* for 10 minutes at 4°C. The aqueous layer was transferred to a new polypropylene tube containing 3 ml of phenol/chloroform, and the samples were vortexed for 1 minute followed by centrifugation as described above. Phenol/chloroform extraction was repeated until no noticeable interface material could be observed, and total nucleic acids were precipitated by adding 1/10 volumes of 3M sodium acetate and 1 volume of isopropanol. The tubes were placed at –70°C overnight, and after centrifugation for 10 minutes at 8740 × *g*, the resulting pellets were washed in 80% ethanol and were redissolved in 450 µl of H_2O. The samples were then transferred to 1.5 ml microcentrifuge tubes and were treated with RNase-free DNAse (Roche, 10 units/µl) to remove chromosomal DNA, by the addition of 1/10 volume of 10× DNAse buffer [0.5 M Tris-HCl, (pH 7.8) and 0.05 M $MgCl_2$] and 7 µl of DNAse. After 30 minutes at room temperature, an additional 7 µl of DNAse was added and the samples were incubated for another 30 minutes. The samples were then extracted with phenol/chloroform followed by an extraction with chloroform, and the RNA was precipitated using 3M sodium acetate and isopropanol as described above. The pellets were washed in 80% ethanol, air-dried and were redissolved in 500 µl of water by placing the tubes in an ice bucket at 4°C overnight. Two 10 µl samples were withdrawn and used for assessing RNA quality and quantity by gel electrophoresis and spectrophotometry, respectively. The RNA preparations were then precipitated as described earlier and were stored at –70°C for future analysis.

2.5.12 Preparation of cell-free extracts

Five-milliliter amounts of *S. clavuligerus* cultures grown in soy and SA media were harvested by centrifugation and were resuspended in 1 ml of lysing buffer [100 mM HEPES (pH 7.2), 0.5 mg/ml lysozyme, 2× Complete EDTA-free protease inhibitor cocktail (Roche)]. The suspensions were incubated at 37°C for 30 minutes and then the mycelia were lysed by sonication, which was carried out on ice using a 2.5 mm probe (Branson Sonifier 450) for 5 × 15 seconds on low setting. The cell debris was removed by centrifugation and the supernatants were aliquoted into 100 µl amounts and were stored at –70°C for future analysis.

The protein concentration in each cell-free extract was determined using the Bio-Rad protein assay (Bio-Rad Laboratories Inc.), which is based on the method described by Bradford (1976).

2.6 Determination of levels of antibiotic production and growth

2.6.1 Growth determination

Growth of *S. clavuligerus* in TSBA was determined by measuring the OD_{600} of the culture. In some cases, a modified version of the method developed by Brana *et al.* (1986) was used to determine the growth of *S. clavuligerus* in fermentation media (Malmberg *et al.*, 1993). Cultures (500 µl) were mixed with 500 µl of 0.62 M HCl and the tubes were sonicated for 30 seconds at setting 1 in a sonication bath. The samples were then diluted in 0.25 M HCl to give OD_{595} values less than 0.6 for determining growth levels (1 OD_{595} unit = 0.59 mg dry cell weight/ml).

During routine fermentation studies, the growth of *S. clavuligerus* was determined using a modified version of a method described by Burton (1957), which indirectly measures deoxyribose. Cultures (500 µl) were microfuged for 10 minutes and the cell pellets were resuspended in 500 µl of 0.4 N perchloric acid (PCA), and were incubated overnight at 4°C. The samples were then microfuged again for 15 minutes at 4°C and the supernatants were discarded. The pellets were resuspended in 500 µl of 0.5 N PCA and the tubes were incubated at 70°C for 2 hours. Then the samples were microfuged at room temperature for 15 minutes and the supernatants were transferred to fresh tubes. Assays were set up in triplicate using FALCON® MICROTEST™ U-Bottom 96 well plates (Becton Dickinson and Co., Franklin Lakes, NJ) by mixing 100 µl of Burton reagent [0.75 g diphenylamine, 50 ml concentrated acetic acid, 0.75 ml concentrated sulfuric acid and 0.25 ml acetaldehyde solution (16 mg/ml)] and 50 µl of the supernatant from above. The plates were then incubated at room temperature from 30 minutes to 24 hours before further analysis. For preparing a standard curve, 0.4 mg of salmon sperm DNA was dissolved in 0.5 M PCA at 70°C for 1 hour. The concentration of the dissolved DNA was confirmed by measuring the A_{260} (A_{260} for a 1 mg/ml solution is 20). The solution was diluted with 0.5 M PCA to give a stock solution containing 0.2 mg/ml salmon sperm DNA. The stock solution was diluted again to give a range between 0.04-0.20 mg/ml DNA, and 50 µl amounts of the diluted standards were used in the assay in triplicate. After the incubation period, the A_{600} of the standards and the samples was measured using a plate reader.

2.6.2 Bioassay for the detection of penicillin and cephamycin

A disc diffusion method using *E. coli* ESS as the indicator organism was used to detect total penicillin/cephamycin production by *S. clavuligerus* (Jensen *et al.*, 1982a). Molten media (150 ml of TSBA) cooled to 50°C, and containing 2 ml of an *E. coli* ESS glycerol stock, was poured into square bioassay plates and allowed to dry for 20-40 minutes. Paper discs were placed on the surface of the agar and 20 µl amounts of *S. clavuligerus* culture supernatants were spotted onto them. The plates were incubated at 37°C overnight before zones of inhibition were measured.

2.6.3 Bioassay for the detection of clavulanic acid

A bioassay using *Klebsiella pneumoniae* ATCC 15380 as the indicator organism was used to detect clavulanic acid in *S. clavuligerus* culture supernatants (Mosher *et al.*, 1999). Square bioassay plates were set up in duplicate and were filled with 150 ml of molten, cooled (50°C) TSBA inoculated with 2 ml glycerol stocks of the indicator organism. For every pair, one of the plates also contained 6 µg/ml Penicillin G. Paper discs were placed on the plates and 20 µl amounts of *S. clavuligerus* culture supernatants were spotted onto them. The zones of inhibition produced after overnight growth at 37°C on the plate containing Penicillin G were attributed to the presence of clavulanic acid in the culture supernatants. They were compared with the zones of inhibition produced by the same samples on the plate without Penicillin G caused by factors other than clavulanic acid.

2.6.4 Bioassay for the detection of alanylclavam

A bioassay using *Bacillus* sp. ATCC 27860 as the indicator organism was used to detect alanylclavam production (Pruess and Kellett, 1983). Square bioassay plates containing 150 ml of Davis minimal medium [0.3% KH_2PO_4, 0.7% K_2HPO_4, 0.05% tri-hydrated sodium citrate, 0.01% $MgSO_4.7H_20$, 0.1% $(NH_4)_2SO_4$, 0.2% glucose and 1.5% agar, adjusted to pH 7] with and without 200 µg/ml L-methionine were used in the assay. The inoculum consisting of indicator organism was prepared by growing the cells on nutrient agar + 2% glucose plates at 28°C for 48 hours. The growth was then scraped into 1 ml of 0.85% NaCl in a microfuge tube and the sample was centrifuged for 1 minute. The pellet was resuspended in 1 ml of 0.85% NaCl and the entire amount was used to inoculate 150 ml of molten, cooled (50°C) Davis minimal medium, before it was poured into plates. After the plates had dried for ~30 minutes, paper discs were placed onto the surface and culture supernatants were spotted on them as described above. After 30-40 hours of growth at 28°C, the zones of inhibition on the plates lacking L-methionine (caused by alanylclavam activity) were measured and compared to those on the plates containing the alanylclavam antagonist, L-methionine, which were caused by other factors independent of alanylclavam production.

2.6.5 HPLC analysis for clavulanic acid and 5*S* clavam production

The quantative determination of clavulanic acid and the 5*S* clavam production by *S. clavuligerus* strains was performed by HPLC analysis of culture supernatants (Foulstone and Reading, 1982; Paradkar and Jensen, 1995). Supernatants were microfuged for 5 minutes to remove mycelia, and 100 µl aliquots of the supernatants were transferred to two sets of fresh tubes. One set of samples was derivatized using 25 µl of 25% imidazole (pH 6.8) while 25 µl of Milli-Q™ water was added to the second set of samples, which functioned as a subtraction control to measure background absorbance. The tubes were incubated at room temperature for 15 minutes, then centrifuged for 5 minutes, and 100 µl amounts of the samples were transferred to polypropylene spring inserts (Fisher) which were placed in HPLC screw neck vials.

The underivatized samples (50 µl) were injected for analysis prior to the derivatized samples using an automated Waters 2690 Separations Module, and a Waters 996 Photodiode Array Detector (Waters, Milford, MA) was used to record the absorbance spectrum. Initially a µBondpack C18 column (Waters, Milford, MA) was used in the analysis, which was replaced by a Phenomenex® Bondclone 10µ C18 (100 × 800 mm 10 micron) column (Phenomenex, Torrance, CA), later on in the study. The samples were analyzed by isocratic elution using a 100 mM $Na_2H_2PO_4$ (adjusted to pH 3.68 with glacial acetic acid) + 6% methanol buffer, at a flow rate of 2 ml/minute, and the runs were 15 minutes each. In the studies conducted using DHA the run times were increased to 30 minutes, and 50 µl blank water injections were used between sample runs to help purge the column of DHA, which was left over from the previous runs. The elution of derivatized clavam compounds was followed by determining the absorbance at 311 nm and the Waters Millenium^{32} Software (version 3.20; Waters, Milford, MA) was used to compile and analyze the data obtained.

2.6.6 LC-MS analysis of clavulanic acid and 5*S* clavam production

Some culture supernatants were also analyzed by LC-MS after derivatization with imidazole as described earlier (Jensen *et al.*, 2004b). Reversed phase liquid chromatography employing an XTerra column (0.21 × 10 cm, Waters Scientific, Milford, MA) at a flow rate of 0.25 ml/minute was used in the analysis. Five microliter amounts of the derivatized samples were injected for HPLC analysis as mentioned in section 2.6.5. In some cases blank water injections were included between samples, to clear the column of material leftover from previous runs. The sample runs were 40 minutes each and the mobile phase consisted of solvent A (10 mM ammonium bicarbonate, pH 10) and solvent B (acetonitrile) used in a binary gradient system as follows: 100% solvent A for 5 minutes, linear gradient to 85% solvent A over 20 minutes, 85% solvent A for 5 minutes, linear gradient to 100% solvent A over 1 minute, 100% solvent A for 9 minutes. The elution of imidazole-derivatized clavam compounds was monitored at 311 nm as described in section 2.6.5, and electrospray mass spectra were acquired on a ZMD-2 single quadrupole instrument (Waters Scientific). The MassLynx software (version 3.4, Waters Scientific) was used to compile and analyze the LC-MS data collected.

2.7 DNA Analysis

Analysis of DNA was carried out using modified versions of procedures described by Kieser *et al.* (2000) and Sambrook *et al.* (1989). The most commonly used methods for analyzing and manipulating DNA isolated from *E. coli* or *S. clavuligerus* are described in the succeeding sections.

2.7.1 Digestion, generation of blunt ended fragments and ligation reactions

Restriction analysis of DNA was carried out as per the manufacturer's recommendations (New England Biolabs, Promega or Roche). For standard subcloning reactions, DNA was digested for 2-4 hours with the appropriate restriction enzymes. For complete digestion, *S. clavuligerus* chromosomal DNA was incubated with the appropriate restriction enzyme for 6 hours to overnight, and an additional aliquot of enzyme was added to the tubes midway through the incubation. Blunt ended fragments were generated using either the Klenow (Roche) or the Mung bean nuclease (New England Biolabs) enzymes. DNA (up to 1 μg) was digested with the appropriate restriction enzyme, which was then heat-inactivated. Nucleotide mix (containing all four dNTPs) was added to the reactions at a final concentration of 0.1 mM along with 2 units of Klenow enzyme, and the tubes were incubated at 37°C. If the Mung Bean Nuclease was used, then DNA (1 μg) was redissolved in 1× Mung Bean Nuclease buffer [50 mM sodium acetate (pH 5.0), 30 mM NaCl and 1 mM $ZnSO_4$] along with 3 units of the enzyme, and the tubes were incubated at 30°C for 30 minutes. The enzymes were inactivated by phenol/chloroform extraction and the DNA was recovered by ethanol precipitation. Overnight ligation reactions were set up at 16°C using T4 DNA Ligase and the corresponding 1× ligation buffers, according to the manufacturer recommendations (New England Biolabs, Roche or U.S. Biochemical). Reactions were 10-30 μl in volumes and the vector to insert ratio was typically kept at 1:2. In ligation reactions involving blunt ended fragments, PEG 8000 (15%) was included in the reaction mixture. The samples were heat-denatured at 50°C for 5 minutes and then chilled on ice for 2 minutes before adding the buffer and 1 μl of the appropriate ligase enzyme. When pCR2.1®TOPO® was used as the cloning vector, PCR products were treated with 1 unit of Taq DNA polymerase (Roche) and 0.2 M dATP in PCR buffer (described in subsection 2.7.3) for 10 minutes at 72°C in order to introduce the necessary 3′ A overhangs, before ligation reactions were set up as per the manufacturer's directions (Invitrogen).

2.7.2 *En bloc* digestion of *S. clavuligerus* DNA for pulsed-field gel electrophoresis

Agarose plugs containing *S. clavuligerus* chromosomal and plasmid DNA were sliced into smaller blocks of approximately 5 mm by 5 mm in dimension. The samples were weighed to determine their volume (1 mg of the gel slice was taken to be 1 μl in volume). The plugs were washed three times in the appropriate 1× reaction buffer in a total volume of 1 ml, 30 minutes each time at room temperature. The entire volume of the gel slice was omitted in determining the amount of reaction buffer required for the third wash, as by this time the gel slice was equilibrated in the 1× reaction buffer. The blocks were placed in the smallest possible volume of 1× reaction buffer, making sure that they were completely submerged. While performing restriction digests, 50-80 units of the corresponding restriction enzymes and 500 μg/ml bovine serum albumin (final concentration, New England Biolabs, Inc.) was incorporated into the reaction. After 4-6 hours of digestion at the appropriate temperature, an additional 10 units of the respective enzyme was added and the blocks were incubated for an additional 6 hours to overnight, before they were subjected to PFGE.

2.7.3 Polymerase chain reaction (PCR)

Polymerase chain reactions were carried out using the EXPAND™ high fidelity PCR system according to the manufacturer's instructions (Roche). Reactions (50 μl) were set up using either 1 μg of genomic DNA or 100-500 ng of plasmid DNA as template. The reactions contained 50 pmol of each primer, 0.2 mM of each dNTP, 5-10% DMSO and 2.5 units of enzyme in 1× Buffer 2 (contains 1.5 mM $MgCl_2$). In reactions where the fidelity of the enzyme was not an issue, 0.5 μl of Taq DNA polymerase or a mixture of Taq and Pfu (1 μl of 1/50 dilution) DNA polymerases (both supplied by The Fermentation Service Unit, Department of Biological Sciences, University of Alberta) was used instead of the EXPAND™ high fidelity enzyme, while all other conditions remained the same. Microfuge tubes (0.5 ml) were used for PCR along with a MiniCycler™ thermocycler (M. J. research, Watertown. MS) after the addition of 2 drops of mineral oil. In some cases, 0.2 ml thin walled tubes (Rose) were used for PCR along with a MiniCycler™ fitted with a heated lid. All oligonucleotide primers used for PCR are listed in Table 2.4.

For the preparation of the REDIRECT© mutagenesis cassettes (Gust *et al.*, 2003), the following program was used along with 5% DMSO: 94°C for 2 minutes followed by 10 cycles of 94°C for 45 seconds, 50°C for 45 seconds and 72°C for 90 seconds, followed by 15 cycles of 94°C for 45 seconds, 55°C for 45 seconds and 72°C for 90 seconds, followed by a final extension step at 72°C for 5 minutes.

Table 2.4: Oligonucleotide primers used in this study.

Primer	Sequence (5′– 3′)[a]	Use[b]
ANW30	TCCTGGCGGGCGAGCTGTACAAC	Forward primer for sequencing the downstream and 3′-end of *cea*S1
ANW31	CGCCTGGGCCTGCCGATCGTC	Forward primer for sequencing the 3′-end of *cea*S1
ANW35	GGGCGGCGAGCAGCGAGGTG	Reverse primer for sequencing the 3′-end of *cea*S1
APA17	GGGGAGTAGCGACAAAACGGTCAGACCCCTGAAGGG	Forward primer for sequence across the deletion in Δ*cea*S2
bls1-S1-For	TCGGATTAATACCTCGCTGCTCGCCGCCCTCAC	Forward primer for *bls*1 S1 probe
bls1-S1-Rev	GGTCGGGGCCGGGCATGGTGAA	Reverse primer for *bls*1 S1 probe
CAN122	GGCCACCGCGACCTGCTGC	Reverse primer for *hrd*B RT-PCR
CAN123	CGGCCAAGCGCACCACTACC	Forward primer for *hrd*B RT-PCR
ceaS1-PR-EX	ATGGCTTTCGCGGTCGTGGT	Reverse primer for *cea*S1 S1 probe and primer extension
ceaS1-RT-For	GCGCAGTCCGAGTCGTAC	Forward primer for *cea*S1 RT-PCR
ceaS1-RT-Rev	TTGGCGGTGTAGGTGGTGAC	Reverse primer for *cea*S1 RT-PCR
ceaS1-S1-For	TCATGAATTCCGGTGGACGGAAGGGGACGG	Forward primer for *cea*S1 S1 probe
ceaS2-PR-EX	GGGCGGTCGATACACGGG	Reverse primer for *cea*S2 S1 probe and primer extension

Table 2.4: (continued).

Primer	Sequence (5′–3′)[a]	Use[b]
ceaS2-FOR	ACGGAGCCTGGTACTGACGGAGTCTGGAGACCGCTCATGATTCCGGGGATCCGTCGACC	Forward primer for *cea*S2 REDIRECT© mutagenesis
ceaS2-REV	ACCCGAAGGCAGCCGGAAGAACCGGTGCCCCCATGATCATGTAGGCTGGAGCTGCTTC	Reverse primer for *cea*S2 REDIRECT© mutagenesis
ceaS2-RT-For	AGGCCGCGTCGATTCTCTTC	Forward primer for *cea*S2 RT-PCR
ceaS2-RT-Rev	CGGCGGGTTGGGGACGGT	Reverse primer for *cea*S2 RT-PCR
ceaS2-S1-For	TGGATCCGTCGCGAATCCAGGGAAGCCGAGC	Forward primer for *cea*S2 S1 probe
cvm7par-FOR	GGCGCCGTCAGCCACGCAGAGAAGATCGGATACGCAGTGATTCCGGGGATCCGTCGACC	Forward primer for *c7p* REDIRECT© mutagenesis
cvm7par-REV	GCGCCGCCCGCCGCCGAGGACCCCGGGGCCCGGGACTCATGTAGGCTGGAGCTGCTTC	Reverse primer for *c7p* REDIRECT© mutagenesis
pah1-S1-For	CCAGATTAATGCGGCGCGGACGGTGCAG	Forward primer for *pah*1 S1 and northern probe
pah1-S1-Rev	CGGGGAGACGGCGGTGGACA	Reverse primer for *pah*1 S1 and northern probe
pah1-UP-Rev	CGCGGCTGCCCCTCCCTC	Reverse primer for *pah*1 S1 probe
KTA1	GGCGAACATGGTCTGGAGCG	Primer to sequence across the *Not*I site of *cea*S2
KTA2	CGCATAATACGACTCACTATAG	Sequencing primer for the T7 end of pWE15

Table 2.4: (continued).

Primer	Sequence (5′– 3′)[a]	Use[b]
KTA3	CGCAATTAACCCTCACTAAAG	Sequencing primer for the T3 end of pWE15
KTA4	TCTCGGCCTCGCGCACCAGCTCAG	Reverse primer to sequence the interior of *cea*S1
KTA5	CATGGCTTTCGCGGTCGTGGTGGC	Reverse primer for sequencing the 5′-end of *cea*S1
KTA6	CGGACGGCGAAGCGGGGTGCG	Reverse primer for sequencing the upstream and 5′-end of *cea*S1
KTA7	CACCGTGCTGACCGCCCTGACCCC	Reverse primer for sequencing the 5′-end of *cea*S1
KTA9	CCTCCTCCTGGCGATCGGCTACG	Reverse primer to sequence the interior of *cea*S1
KTA10	AGCCCCGTGACATCGCCGTAGAC	Forward primer for sequencing the upstream and 5′-end of *cea*S1
KTA11	CCACCACGACCGCGAAAGCCATG	Forward primer for sequencing the 5′-end of *cea*S1
KTA12	CGAGCTGAACGGCACCCCCACC	Forward primer to sequence the interior of *cea*S1
KTA13	CGCCGTCAACCGCTATCCGTACTC C	Reverse primer for sequencing the upstream region of *cea*S1

Table 2.4: (continued).

Primer	Sequence (5′– 3′)[a]	Use[b]
KTA14	CCATCCCGGCGCCCGTCCGATGCGAAGGAGATCTCCATGATTCCGGGGATCCGTCGACC	Forward primer for *cea*S1 REDIRECT© mutagenesis
KTA15	CGGGGCCGGGCATGGTGAACTCGTCCTCCACGGTGGTCATGTAGGCTGGAGCTGCTT	Reverse primer for *cea*S1 REDIRECT© mutagenesis
KTA-ceaS1-For	GCG**GGATCC**GGGCGGTCAGCACGGT	Forward primer for cloning the *cea*S1 promoter
KTA-ceaS1-Rev	CCG**GGTACC**AGGGTCGCGAAGCACG	Reverse primer for cloning the *cea*S1 promoter
KTA-ceaS2-For	AACCCCA**GGATCC**GAGCCCCACCGTCACG	Forward primer for cloning the *cea*S2 promoter
KTA-ceaS2-Rev	CGGCCG**GGTACC**CCAAACACCTTCCCCACAC	Reverse primer for cloning the *cea*S2 promoter
KTA-GFP-REV	CCGGTGAACAGCTCCTCG	Reverse primer complementary to *egfp* for sequencing upstream regions in promoter probe vectors
PBE16	CGTTCCTGATCGAGGTACCGG	Forward primer for sequencing the 3′-end and downstream region of *cea*S1
PBE27	CTCGGCGCGGTTGTACAGCTCG	Reverse primer for sequencing the 3′-end and downstream region of *cea*S1

Table 2.4: (continued).

Primer	Sequence (5′– 3′)[a]	Use[b]
Red-SEQ-UP	CTGCAGGTCGACGGATCC	Reverse primer for sequencing the region upstream of the Redirect disruption cassette
Red-SEQ-DWN	CGAAGCAGCTCCAGCCTAC	Forward primer for sequencing the region downstream of the Redirect disruption cassette
sen-kin-FOR	GCGTCATCGGGCGGGTAGGGGGGCCTGGCAGCGGTCGTGATTCCGGGGATCCGTCGACC	Forward primer for *skn* REDIRECT© mutagenesis
sen-kin-REV	CGACCGCCGGGGCGGTCCCGGGTGCGGCGGCCGGATCTATGTAGGCTGGAGCTGCTTC	Reverse primer for *skn* REDIRECT© mutagenesis

[a] Non homologous sequences incorporated into oligonucleotide primers are underlined and engineered restriction sites are shown in bold letters

[b] Details of the function of each oligonucleotide primer are described under Materials and Methods and Results

All double stranded DNA probes for S1 nuclease protection assays were prepared by PCR using custom primers. DMSO (5% v/v) was included in the reactions with the following program: 94°C for 2 minutes followed by 10 cycles of 94°C for 45 seconds, 60°C for 45 seconds and 72°C for 45 seconds, followed by 15 cycles of 94°C for 45 seconds, 65°C for 45 seconds and 72°C for 45 seconds, with a final extension step at 72°C for 5 minutes.

The *cea*S1 and *cea*S2 promoter regions were isolated as PCR fragments for subcloning and subsequent expression analysis using the same general two-step PCR programs mentioned above, except that the annealing temperatures were 55°C and 65°C, respectively.

2.7.4 Agarose gel electrophoresis

Electrophoresis was carried out using 0.8-1.0% gels in TAE buffer (40 mM Tris-acetate and 1 mM EDTA). Lambda DNA (Roche) digested with *Pst*I or *Bst*EII was used as the molecular weight marker. Loading buffer (5×) [60% sucrose, 100 mM EDTA (pH 8), 0.25% bromophenol blue, 0.25% ficoll orange and 0.25% xylene cyanol] (Hopwood *et al.*, 1985) was used to increase the density of the DNA sample and to visualize the migration of dyes in the gels. To visualize RT-PCR products 1.5% agarose TBE (90 mM Tris, 89 mM boric acid and 2.5 mM EDTA, sodium salt) gels were used. All gels were stained in 2.0 μg/ml ethidium bromide in Milli-Q™ water for 5-10 minutes before they were photographed under UV light.

2.7.5 Pulsed field gel electrophoresis

After *en bloc* digestion of *S. clavuligerus* DNA, the blocks were removed from the tubes and placed into the wells of a 1.5% agarose gel containing 12 mM thiourea. The wells were sealed using 1.5% low melting agarose in HE buffer and the gel was placed in the running buffer for equilibration for 1 hour. Pulsed-field gel electrophoresis was performed using a BIO-RAD CHEF-DR® contour clamped homogenous electric field apparatus using 0.5× TBE buffer + 50 μM thiourea (Evans and Dyson, 1993). Gels were run at 14°C and 170 volts for 45 hours with a ramp time of 10-150 seconds. Yeast chromosome PFG markers and MidRange-I PFG markers (New England Biolabs, Inc.) were run alongside *S. clavuligerus* chromosomal DNA for size estimation. After the runs were complete, the DNA was stained for exactly 30 minutes in 1.0 μg/ml ethidium bromide and visualized as described in section 2.7.4.

2.7.6 Polyacrylamide gel electrophoresis (PAGE)

Where appropriate, small DNA fragments were fractionated using 8% polyacrylamide (29:1 acrylamide:N,N′-methylene bisacrylamide, Bio-Rad) TBE gels. DNA was electrophoresed at 100 V using lambda DNA markers and 5×-loading dye as described above for agarose gel electrophoresis. The gels were stained with ethidium bromide (1.0 μg/ml) for 30 minutes before visualization under UV light.

2.7.7 Gel purification of DNA

DNA fragments fractionated by agarose gel electrophoresis were purified using the QIAquick Gel Extraction Kit as per the manufacturer's directions (Qiagen Inc.). DNA was eluted from the spin columns using either 30 μl of sterile Milli-Q™ water or the elution buffer provided with the kit, by incubating the tubes at room temperature for 5 minutes and at 50°C for 2 minutes, before the final spin. DNA fragments fractionated by PAGE were isolated using the crush and soak method (Sambrook *et al.*, 1989). Briefly, DNA fragments were excised and placed into a 0.5 ml microfuge tubes having a hole poked at its base with a hot, large gauged needle. The 0.5 ml microfuge tube was placed in a larger 1.5 ml tube and the sample was spun for 2 minutes to crush the gel slice by squeezing it through the fine hole, into the larger 1.5 ml microfuge tube. The sample was weighed (1 mg of the gel slice was taken to be 1 μl) and 2 volumes of elution buffer [0.5 M sodium acetate, 10 mM magnesium acetate, 1 mM EDTA (pH 8) and 1% SDS] was added to the tube, which was then incubated at 37°C on a drum roller for 6-16 hours. During the incubation, the gel slices were homogenized once more using a small pestle. The tube was then microfuged and the supernatant was transferred to a fresh tube. The residual pellet was resuspended in 0.5-1 volumes of the elution buffer by vortexing, and the tube was microfuged again. The supernatants from both fractions were combined and microfuged once more to remove any residual polyacrylamide. The supernatant was transferred to a fresh tube, extracted with phenol/chloroform and the DNA was precipitated by ethanol precipitation along with the addition of 1 µl of glycogen (20 mg/ml, Roche).

2.7.8 Labeling of DNA probes

Double stranded DNA fragments used as probes in Southern and northern hybridizations were labeled by nick translation (Hopwood *et al.*, 1985; Sambrook *et al.*, 1989). Reactions (30 µl) contained 100-500 ng of DNA, 1µl each of 1 mM unlabelled dATP, dGTP and dTTP and a 10 µCi amount of [α-^{32}P]dCTP (Amersham) in 1× NT (nick translation) buffer [50 mM Tris (pH 7.4), 10 mM $MgCl_2$, 1 mM DTT (Dithiothreitol) and 50 µg/ml BSA]. DNAse I (2.5 µl of a 10^{-4} dilution, Roche) and 0.5 µl of DNA polymerase I (2.5 units, Roche) were then added and the tubes were incubated at 16°C for 1-2 hours. Reactions were terminated by adding 15 µl of 0.5 M EDTA (pH 8). The DNA probes were stored at –20°C or they were denatured at 90°C for 5 minutes, chilled on ice for 2 minutes and used immediately in hybridization studies.

Double stranded DNA probes for 5′ S1 nuclease protection analysis and single stranded DNA oligonucleotides used for primer extension analysis were labeled at the 5′ end using a modified version of the end labeling procedure described by Chaconas and van de Sande (1980). Approximately 100 ng of double stranded DNA or 50 pmol of oligonucleotide primers were used in 10 µl final reaction volumes with 1× kinase reaction buffer [50 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 5 mM DTT and 0.1 mM spermidine], and 5 µl (50 µCi) of γ-^{32}P-ATP (ICN). T4 polynucleotide kinase (10 units/µl, Roche) was diluted 1/10 in kinase dilution buffer [50 mM Tris-HCl (pH 7.6), 1 mM DTT, 0.1 mM EDTA and 50% glycerol] and 1 µl of the diluted enzyme was added to the reaction mixture. The tubes were then incubated at 37°C for 15 minutes, after which a second aliquot (1 µl) of the diluted enzyme was added and the reactions were incubated for another 15 minutes. Sterile water (40 µl) was then added to the reaction mixtures and the DNA was ethanol precipitated along with the addition of 1 µl of glycogen (20 mg/ml). The samples were left on ice for 2-3 hours, and then the DNA was pelleted by microfugation and washed in 80% ethanol before it was redissolved in an appropriate volume of RNase-free water.

2.7.9 Southern blotting

Southern blotting was performed according to a modified version (Sambrook *et al.*, 1989) of the original procedure described by Southern (1975). DNA was fractionated by conventional agarose gel electrophoresis or by PFGE, stained and visualized as described in sections 2.7.4 and 2.7.5, respectively. Conventional gels were depurinated by placing them in 0.25 M HCl, for 10-20 minutes with gentle agitation, or until the bromophenol dye present in the gels changed color. DNA fractionated by PFGE was nicked using a BIO-RAD GS GENE LINKER™ UV CHAMBER (60 mJoule), instead of depurination treatment. DNA was denatured by placing the gels in denaturation buffer (1.5 M NaCl and 0.5 M NaOH) for 30 minutes with gentle agitation, or until the bromophenol blue dye reverted to its original color for conventional gels. The gels were then subjected to neutralization treatment by washing in 3 M NaCl and 0.5 M Tris-HCl (pH 7.5) for 30 minutes.

Capillary transfer under neutral conditions was used to transfer fractionated DNA fragments from gels onto Hybond™-N nylon membranes as described by Sambrook *et al.* (1989). Gels were placed face down onto a wick prepared by combining three strips of chromatography paper, soaked in 20× SSC (3 M NaCl and 0.3 M tri-sodium citrate). A piece of membrane similar in dimension to the gel was placed on top of each gel followed by three pieces of chromatography paper, also cut to the same dimensions and pre-wet in 20× SSC. A stack of paper towels ~15 cm in height was placed on top of each gel sandwich. A glass dish with a bottle containing 2 L of water was then placed on top of each setup to weigh it down. The transfer was allowed to continue overnight, after which the apparatus was disassembled and the locations of the wells were marked on the membranes. The DNA was fixed by baking the membrane as described in section 2.4.6, and the membrane was either used immediately or stored at room temperature, sandwiched between two pieces of dry chromatography paper, wrapped in a plastic bag for future use.

2.7.10 Southern and colony hybridization

Hybridizations were carried out in hybridization bottles (Robbins Scientific) using a H-12000 hybridization incubator (Tyler Research Instruments Corp. Edmonton, AB). Membranes were placed into the bottles and were washed once in 2× SSC, which was then replaced with pre-hybridization solution consisting of the hybridization buffer [5× SSC, 5× Denhardt's solution (2% Ficoll™ 400, 2% polyvinylpyrolidone and 20 mg/ml BSA) and 0.5% SDS] (Denhardt, 1966), and 100 μg/ml denatured salmon sperm DNA. The membranes were allowed to wash in the pre-hybridizing solution for 1-2 hours at 65°C, after which the pre-hybridizing solution was replaced with 15 ml of fresh hybridization buffer. Denatured, labeled probe was then added to the bottles, which were incubated overnight at 65°C. After hybridization, the membranes were washed twice for 15 minutes in 2× SSC + 0.1% SDS at 21°C, and then once for 20 minutes in 1× SSC + 0.1% SDS, and once for 20 minutes in 0.1× SSC + 0.1% SDS, all at 65°C.

If the membranes needed to be re-hybridized, they were first stripped of bound radioactivity by treating twice with 0.2 M NaOH at 42°C for 15 minutes. The blots were then washed twice with 2× SSC at 65°C for 20 minutes each, and removal of the bound probes was confirmed using a phosphorimager.

2.7.11 DNA sequencing

DNA sequencing was carried out using the dideoxynucleotide chain terminator method (Sanger *et al.*, 1977). Automated DNA sequencing was performed using either the DYEnamic ET Terminator (Amersham Pharmacia) or the Big Dye v3.1 (Applied Biosystems), Cycle Sequencing Kits. Reactions (20 μl) were set up using 300-500 ng of plasmid DNA or 600-800 ng of cosmid DNA and 5-10 pmol of primer, all dissolved in sterile Milli-Q™ water. DNA for automated sequencing was quantified using a TKO 100 DNA fluorometer (Hoefer Scientific Instruments, San Francisco, CA). For the ET kit, 8 μl amounts of the reaction mix were routinely used, whereas Big Dye reactions were set up using 2 μl amounts of the premix along with 6 μl of the dilution buffer [200 mM Tris (pH 9.0) and 5 mM $MgCl_2$]. DMSO (5%) was routinely included in the reactions with the program consisting of 25-30 cycles of denaturation at 96°C for 30 seconds, annealing

(~5°C below the Tm of the primer being used) for 15 seconds and extension at 60°C for 2 minutes. Reactions were performed in a MiniCycler™ thermocycler (M. J. research, Watertown, MS) using 0.6 ml microfuge tubes and two drops of mineral oil was layered on top of the aqueous phase to minimize evaporation. Once the reaction cycle was completed, the lower fraction was transferred to a new microfuge tube and 2 µl of a salt solution [1.5 M sodium acetate (pH>8.0) and 250 mM EDTA] and 80 µl of 95% ethanol was added to each tube to precipitate the reaction products. The tubes were placed on ice for 15 minutes followed by microfugation for 15 minutes. The pellets were washed with 500 µl of 70% ethanol, air dried and provided to the Molecular Biology Service Unit, University of Alberta, where the sequence information was obtained.

Manual sequencing of DNA was performed using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit as per the manufacturer's directions (USB). Reactions containing 300-500 ng of template were set up with 20 mM Tris-HCl, (pH 9.5), 6 mM $MgCl_2$, 2 pmol of oligonucleotide primer, 2 µM dNTPs, 0.02 µM ^{33}P-ddNTPs and 8 units Thermo Sequenase™ enzyme. The following program was used for manual DNA sequencing: 95°C for 2 minutes followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing (~5°C below the Tm of the primer being used) for 30 seconds and extension at 72°C for 1 minute. Reactions were carried out using 0.2 µl thin-walled tubes and a Perkin Elmer GeneAmp PCR System 2400 Thermocycler. Loading dye [98% deionized formamide, 10 mM EDTA (pH 8), 0.025% bromophenol blue and 0.025% xylene cyanol] was then added (4 µl), and 3-5 µl of each reaction was heated for 5 minutes at 95°C and loaded onto a 6% polyacrylamide sequencing gel [19:1 acrylamide:N,N'-bisacrylamide (Bio-Rad), 8.3M urea and 1× TBE]. The samples were electrophoresed at 40 watts for 1.75-3 hours, and the gel was transferred to 3MM Whatman No.1 filter paper and dried for 2 hours at 80°C on a Bio-Rad gel dryer connected to a Savant pump. The dried gel was then exposed overnight to a phosphorscreen.

2.8 RNA Analysis

2.8.1 Northern analysis

Northern analysis was performed as described previously (Williams and Mason, 1985), using 40 μg RNA isolated from wild type *S. clavuligerus* grown in soy medium for 96 and 120 hours. RNA samples and 1 μl molecular weight marker (Marker III, Roche) were dissolved in a final volume of 2.5 μl DEPC (Diethyl pyrocarbonate)-treated Milli-Q™ H_2O. The samples were then incubated at 50°C in a glyoxal/DMSO solution [1 M glyoxyl, 50% DMSO and 10 mM Na_2HPO_4/NaH_2PO_4 (pH 7)] in order to denature the nucleic acids. RNA loading dye [50% glycerol, 10 mM Na_2HPO_4/NaH_2PO_4 (pH 7), 0.4% bromophenol blue and 0.4% xylene cyanol) was added to each sample, which were then loaded onto a 1.25% Na_2HPO_4/NaH_2PO_4 agarose gel. Electrophoresis was carried out at 60V until the bromophenol blue indicator dye had migrated 2/3 the length of the gel. The samples were transferred onto a Hybond™-N membrane overnight using capillary transfer. After marking the positions of the wells, the RNA was UV cross-linked to the membrane, which was then baked for 1 hour under vacuum at 80 °C in order to remove the glyoxyl. The membrane was either used immediately for northern hybridization or was wrapped in saran wrap and stored at –20°C.

2.8.2 High resolution S1 nuclease protection analysis

High-resolution S1 nuclease mapping was carried out using the sodium trichloroacetate (NaTCA) method (Kieser *et al.*, 2000). All double stranded DNA probes (see 2.7.3) were prepared by PCR using custom primers (Table 2.4). RNA (40 μg, stored as an isopropanol precipitate) was pelleted and washed in 80% ethanol, and after removal of all the residual ethanol, the pellets were redissolved in 1 μl of the end labeled probe (dissolved at a concentration of 30 fmol/μl in RNase-free water). The samples were then dried in a Speed Vac (Savant) for approximately 2 minutes. Twenty microliters of NaTCA hybridization buffer (3M NaTCA, 50 mM PIPES and 5 mM EDTA, adjusted to pH 7.0 with 5 M NaOH) was then added to the tubes. To help dissolve the pellets, the tubes were heated at 70°C for 2 minutes followed by vortexing, until all the contents had gone into solution. The tubes were placed at 70°C for 15 minutes and were immediately transferred to a water bath set at 45°C. After overnight incubation, one tube was removed at a time, briefly centrifuged, and 300 μl of S1 digestion mix [0.28 M NaCl, 15 mM sodium acetate (pH 4.4), 4.5 mM zinc acetate, 20 µg/ml partially-cleaved denatured calf thymus DNA and 100 units of S1 nuclease (Roche)], was added to each tube. The tubes were immediately vortexed and placed on ice, and after all the tubes had been processed, they were briefly centrifuged and placed at 37°C. After 45 minutes of incubation at 37°C, 75 µl of S1 termination solution (2.5 M ammonium acetate and 0.05 M EDTA) was added to each tube, after which they were vortexed and then microfuged briefly. The DNA/RNA hybrid complexes were precipitated by adding 2.5 µl of glycogen (20 mg/ml, Roche) and 400 µl of isopropanol, and placing the tubes on ice for 2-3 hours followed by centrifugation. The supernatant was removed carefully and the pellets were dissolved in 3 µl of sequencing loading dye (as described in section 2.7.11). Samples (1-2 µl) were run on 6% polyacrylamide sequencing gels as described in section 2.7.11. The template plasmids and the reverse primers used to prepare the S1 probes by PCR were used in sequencing reactions, which were run alongside the S1 protection reaction products for size estimation.

2.8.3 Primer extension analysis

Primer extension analysis was performed using the *C. therm* polymerase for reverse transcription in two-step RT-PCR according to the manufacturer's instructions (Roche), with the following changes. Twenty microliter reactions were set up using 5-pmol of the end labeled reverse primers and 40 units of RNaseOUT™ Recombinant Ribonuclease Inhibitor (Invitrogen). The following program was typically used in the analysis: extension at 60°C for 60 minutes and termination at 80°C for 10 minutes. Sequencing ladders were prepared for size estimation using the same reverse primers used for primer extension analysis and template plasmids encoding the 5′ and upstream regions of genes, for which the TSPs were being determined. The sequencing ladders and the samples were separated on 6% denaturing polyacrylamide-sequencing gels for analysis as described in section 2.7.11.

2.8.4 RT-PCR analysis

Reverse transcriptase polymerase chain reaction (RT-PCR) analysis of RNA was carried out using the *C. therm* polymerase for reverse transcription in two-step RT-PCR

(Roche). All RT reactions were carried out at 62°C for 30 minutes as per the manufacturer's instructions with the following changes. The reactions were set up in a final volume of 10 µl using 0.5 µg of total RNA per reaction and 15.8 units of RNAguard Ribonuclease Inhibitor (Amersham). The PCR reactions were performed using 5 µl of the RT product from above in a final volume of 50 µl and the EXPAND™ high fidelity PCR system with Buffer 2. In addition, control reactions were also set up using the same RNA preparations without subjecting them to the RT step to verify that the results observed were not due to chromosomal DNA contamination. DMSO (5% v/v final concentration) was used for the PCR amplification of the *cea*S1 and the *cea*S2 RT-products with the following program: 94°C for 2 minutes followed by 25 cycles of 94°C for 1 minute, 62°C for 1 minute and 72°C for 1 minute. The *hrd*B RT-product was amplified by PCR using 10% v/v DMSO and the program: 94°C for 2 minutes followed by 25 cycles of 94°C for 1 minute, 66°C for 1 minute and 72°C for 1 minute. The identities of the RT-PCR products were also verified by restriction analysis followed by PAGE and by sequencing of both DNA strands.

2.9 Protein analysis

2.9.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

S. clavuligerus cell-free extracts were analyzed by denaturing SDS-PAGE as described earlier (Blackshear, 1984; Sambrook *et al.*, 1989). Samples were electrophoresed on a 12% SDS-PAG using a Bio-Rad Mini Protean cell 3 apparatus. The resolving gel consisted of 12% polyacrylamide (29:1 acrylamide: N,N′-methylene bisacrylamide), 0.2% NaCl, 375 mM Tris-HCl (pH 8.8) and 0.1% SDS, and the stacking gel consisted of 4.0% polyacrylamide (29:1 acrylamide: N,N′-methylene bisacrylamide), 125 mM Tris-HCl, pH 6.8 and 0.07% SDS. The gels were run using freshly prepared SDS-PAGE running buffer [50 mM Tris-HCl (pH 8.3), 384 mM glycine and 0.1% SDS]. Kaleidoscope pre-stained standards (Bio-Rad) were used as the molecular weight markers. Samples were prepared for electrophoresis by adding cell extract to an equal volume of 2× loading buffer [125 mM Tris-HCl (pH 6.8), 4% SDS, 10% 2-mercaptoethanol, 0.5% bromophenol blue and 20% glycerol] and boiling for 5 minutes.

Cell-free proteins (50 μg) were separated by running gels at 50-200 V at either 4°C or at room temperature.

2.9.2 Western analysis

After SDS-PAGE, *S. clavuligerus* cell-free extracts and molecular weight markers were transferred to PVDF membranes (Stratagene, La Jolla, CA) using a Bio-Rad Transblot apparatus according to the method of Towbin *et al.* (1979). PVDF membranes were pre-wetted by first dipping in methanol and then by equilibrating them for 10 minutes in western transfer buffer [25 mM Tris-HCl (pH 8.3), 150 mM glycine and 20% methanol]. The gel containing fractionated proteins was also soaked in western transfer buffer for 15 minutes and then placed on the pre-wet PVDF membrane. The gel with the membrane in direct contact was bound on both sides with pieces of chromatography paper, cut to match the dimensions of the gel, followed by pieces of foam, all of which had been presoaked in the transfer buffer. The transfer was done for 1-2 hours at 200 mAMPS and 4°C.

The BM Chemiluminescence Western Blotting Kit (Mouse/Rabbit) was used as the secondary antibody to detect proteins as per the manufacturer's instructions (Roche). The primary antibody used to detect EGFP consisted of the commercially available BD Living Colors A.v. Peptide Antibody (BD Biosciences), and was used at 1/400 dilution. All washes and subsequent procedures were carried out using the materials and instructions provided along with the secondary antibody detection kit.

2.10 Confocal Microscopy

S. clavuligerus cultures harboring *egfp* reporter constructs were grown in TSBS for 36 hours. Five hundred microliter amounts of the TSBS cultures were used to inoculate 25 ml of either SA or soy medium, and after 72 hours of growth 1 ml amounts of the cultures were harvested and washed once in acetonitrile and then twice in sterile distilled water. The washed mycelia were suspended in 500-1000 μl of 40% glycerol and 10 μl amounts were mounted under No. 0 Microscope cover slips (CANEMCO Supplies, St. Laurent, Quebec). The edges of the cover slips were sealed using clear nail polish before observations were taken under the microscope.

Confocal microscopy was carried out using a Leica DM IRB Inverted microscope. An argon laser (50-52% attenuation) provided excitation at 488 nm. Fluorescence due to EGFP excitation was detected between 505 nm to 520 nm and corresponding Differential Interference Contrast (DIC) images were also obtained. Images were collected using the Leica confocal microscopy software and were processed using Adobe®Photoshop 7.0®.

2.11 Software and web sources used in analysis

DNA sequence information obtained was compiled using either DNA Strider 1.2 (Marck, 1988) or GeneTools 1.0 (BioTools Incorporated). Open reading frames (ORFs) were predicted based on codon preference using the online program FramePlot 3.0beta (Bibb *et al.*, 1984; Ishikawa and Hotta, 1999) at (http://watson.nih.go.jp/~jun/cgi-bin/frameplot-3.0b.pl) and similarity searches were performed using the online BLAST program (Altschul *et al.*, 1990; Altschul *et al.*, 1997; Gish and States, 1993) at the National Center for Biotechnology Information (http://www.ncbi.nih.gov/BLAST/). The PROSITE online program at the ExPASy home page was used to search for specific peptide motifs (http://www.expasy.org/prosite/).

Three dimensional protein structures were predicted using the SWISS-MODEL (Guex and Peitsch, 1997; Peitsch, 1995; Schwede *et al.*, 2003) prediction server (http://swissmodel.expasy.org/) and were viewed using the Deep View Swiss PBD Viewer program (http://swissmodel.expasy.org/). Potential RNA secondary structures, together with ΔG values, were determined using the Mfold program (Version 3.1) (Mathews *et al.*, 1999; Zuker, 2003) available online at (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi).

CHAPTER 3: RESULTS

3. Results

The *cea*S gene, encoding carboxyethylarginine synthase (CeaS), was independently isolated and described by two separate groups, and was found to be located in the clavulanic acid gene cluster of *S. clavuligerus* (Jensen *et al.*, 2000; Perez-Redondo *et al.*, 1999)(Figure 1.4.2). Studies to be described later in this thesis will show that there are in fact two copies of the *cea*S gene, present in *S. clavuligerus*. Therefore, for clarity, the original *cea*S gene is here forth referred to as *cea*S2 due to its proximity to *cas*2, from the clavulanic acid gene cluster (Jensen *et al.*, 2000; Perez-Redondo *et al.*, 1999)(Figure 1.4.2). When *cea*S2 was initially isolated, its function was unknown, and its predicted gene product shared approximately 29% identity with acetohydroxy or acetolactate synthase (AHAS) large subunits from different bacteria (Perez-Redondo *et al.*, 1999).

To investigate the involvement of *cea*S2 in clavulanic acid and 5*S* clavam biosynthesis, Pérez-Redondo *et al.* (1999) and Elder (1998) [note that portions of K. Elder's work was published as part of a study by Jensen *et al.*, (2000)] prepared *S. clavuligerus cea*S2 mutant strains and tested them for their ability to produce the respective metabolites. Both the above-mentioned studies showed that *cea*S2 did not encode an active AHAS, and that it was somehow involved in clavulanic acid and 5*S* clavam biosynthesis. It was found that the *cea*S2 mutants did not produce any clavulanic acid or 5*S* clavams when grown on either SA (Jensen *et al.*, 2000) or SPG (GSPG without glycerol) (Perez-Redondo *et al.*, 1999) medium. However, when the *cea*S2 mutants were grown on either soy (Jensen *et al.*, 2000) or GSPG (Perez-Redondo *et al.*, 1999) fermentation medium, some clavulanic acid and 5*S* clavam production was still observed. It was subsequently demonstrated by Khaleeli *et al.* (1999) that CeaS2 catalyzes the condensation of a molecule of L-arginine and glyceraldehyde-3-phosphate to give N^2-(2-carboxyethyl)arginine (Figure 1.4.1), the first reaction in the shared clavulanic acid and 5*S* clavam biosynthetic pathways (Elson *et al.*, 1993a).

Since CeaS2 catalyzes a unique reaction (Khaleeli *et al.*, 1999), it seemed most likely that the observed production of clavulanic acid and the 5*S* clavams by the *cea*S2 mutant in soy and GSPG media, was due to the presence of another functionally similar enzyme, and not due to some sort of nonspecific activity. In the case of CAS (Marsh *et* *al.*, 1992; Salowe *et al.*, 1990) and PAH (Jensen *et al.*, 2004b; Wu *et al.*, 1995), it was known that there were two copies of paralogous genes (*cas*1/*cas*2 and *pah*1/*pah*2, respectively) encoding two isozymic forms of the respective proteins in *S. clavuligerus*. Based on similarities in the observed phenotypes of the *cea*S2, *pah*2 and *cas*2 mutants (Table 1.4.1), it was proposed that a paralogue might also exist for *cea*S2 (Jensen *et al.*, 2000; Perez-Redondo *et al.*, 1999). It was also suggested that the two paralogous *cea*S genes might be differently regulated in terms of nutrition, depending on the fermentation medium used (Jensen *et al.*, 2000).

The main objective of this study was to isolate the second, paralogous copy of the gene encoding CeaS in *S. clavuligerus*. The roles of the respective paralogous genes in clavulanic acid and 5*S* clavam metabolite biosynthesis, and the mechanisms involved in regulating their transcription, were also examined. In addition, the physical linkage or the location of the gene clusters on the *S. clavuligerus* chromosome, which comprise the genes involved in clavulanic acid and 5*S* clavam metabolite biosynthesis, was investigated. This was done to determine the relationship between the respective gene clusters, and to provide further insight into the lineage of the paralogous pairs of genes, involved in the early stages of clavulanic acid and 5*S* clavam metabolite biosynthesis in *S. clavuligerus*.

3.1 Preparation and analysis of new *S. clavuligerus cea*S2 mutants

In a previous study, where a *cea*S2 mutant strain was prepared by insertion of an apramycin resistance gene cassette into the gene, in the direction opposite to *cea*S2 transcription (*cea*S2*::apr*)(Elder, 1998), the mutant strain was found to be severely compromised in its ability to produce clavulanic acid and the 5*S* clavams (Table 1.2). The respective metabolites were produced at such low levels that they were only detected using very sensitive bioassays and not by HPLC analysis of culture supernatants. Due to the nature and importance of the reaction catalyzed by CeaS in the clavulanic acid and 5*S* clavam biosynthetic pathways (Figure 1.4.1), the production of the respective metabolites by the *cea*S2 mutant was hard to explain, without proposing the existence of a second copy of the *cea*S gene in *S. clavuligerus* (Jensen *et al.*, 2000). At the time it was known that *cea*S2, *bls*2, *pah*2 and *cas*2 from the clavulanic acid gene cluster were transcribed as part of a 5.3 kb polycistronic transcript (Paradkar and Jensen, 1995)(Figure 1.4.2), and it was proposed that the *cea*S2*::apr* mutation could have polar effects on the expression of the downstream *bls*2, *pah*2 and *cas*2 genes, due to the insertion of the *apr* gene cassette, which could explain the very severe phenotype of the mutant. Therefore, new *cea*S2 mutants were prepared in order to overcome the effects of polarity on the expression of genes downstream of *cea*S2. This was also done to provide additional evidence for the possibility of the presence of a second copy of the *cea*S gene in *S. clavuligerus*, which could functionally replace the *cea*S2 gene product, provided there were no polar effects associated with the *cea*S2 mutation. Note: portions of this section have been published as part of the manuscript by Tahlan *et al.* (2004c).

3.1.1 Preparation and analysis of the *cea*S2 frame shift (*cea*S2-Fs) mutant

A new *cea*S2 mutant was prepared in which the disrupting *apr* gene cassette present in the *cea*S2*::apr* mutant (Jensen *et al.*, 2000), was replaced by a simple frame shift mutation, to alleviate potential polar effects. This gave rise to the *cea*S2 frame shift (*cea*S2-Fs) mutant.

3.1.1.1 Preparation of the *cea*S2-Fs mutant

The strategy used for the preparation of the *cea*S2-Fs mutants using the plasmid pCAD2-3 has been outlined in Figures 3.1.1 and 3.1.2. The plasmid pCAD2-3 was linearized by digestion at a unique *Not*I restriction endonuclease site, located 674 bp from the proposed start codon of *cea*S2. The linearized plasmid was made blunt by treatment with the Klenow fragment of DNA polymerase I, and was then re-circularized to give pCAD2-3(L1-5), which had a 4 bp insertion resulting in a +1 frame shift mutation in *cea*S2 (Figure 3.1.1). The insertion of the frame shift mutation in *cea*S2 was confirmed by DNA sequencing using the primer KTA1, and by using restriction analysis to verify that the *Not*I site in pCAD2-3(L1-5) was destroyed by the insertion of the four nucleotides. The *Streptomyces* vector pIJ486, which is a derivative of pIJ101 and is easily lost from *Streptomyces*, was fused to pCAD2-3(L1-5) at the *Hin*dIII site to give the *E. coli-Streptomyces* shuttle vector pCAD2-3(L1-5)486. The plasmid pCAD2-3(L1-5)486

Figure 3.1.1 Introduction of the frame shift mutation in *cea*S2. The sequence of the *Not*I restriction site in *cea*S2 is shown and the rest of the gene present in the plasmid pCAD2-3 is represented by the thick lines. The bold letters represent the extra nucleotides incorporated into *cea*S2, resulting in the frame shift mutation.

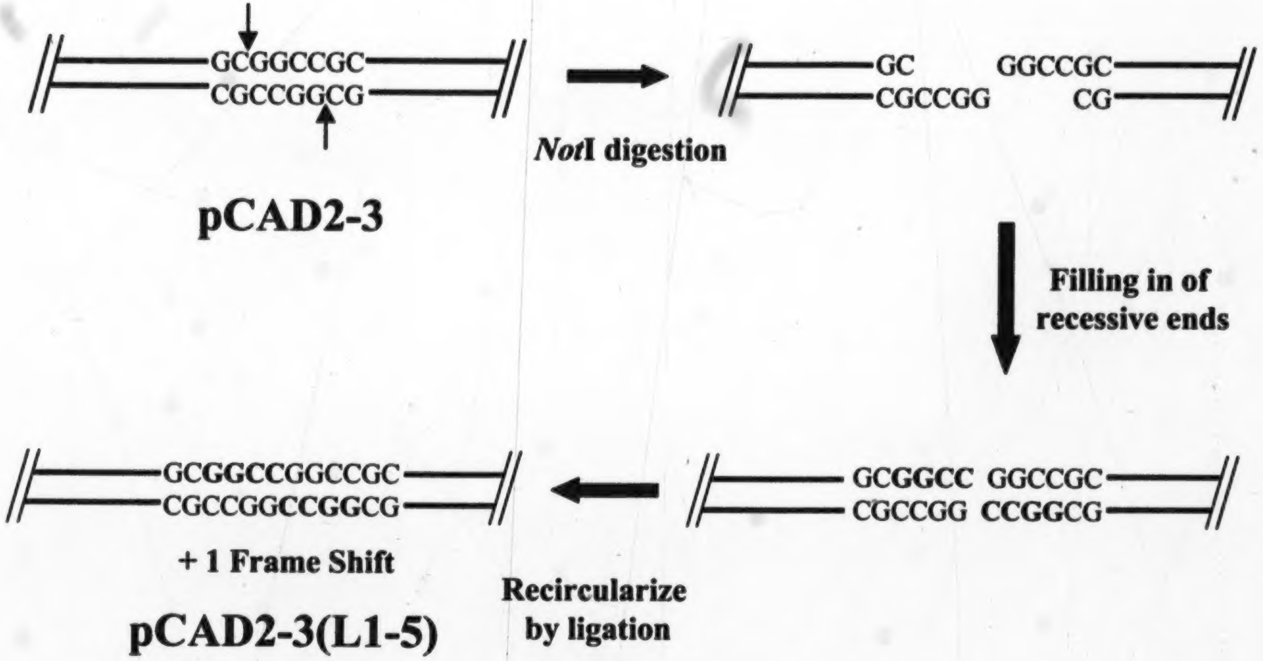
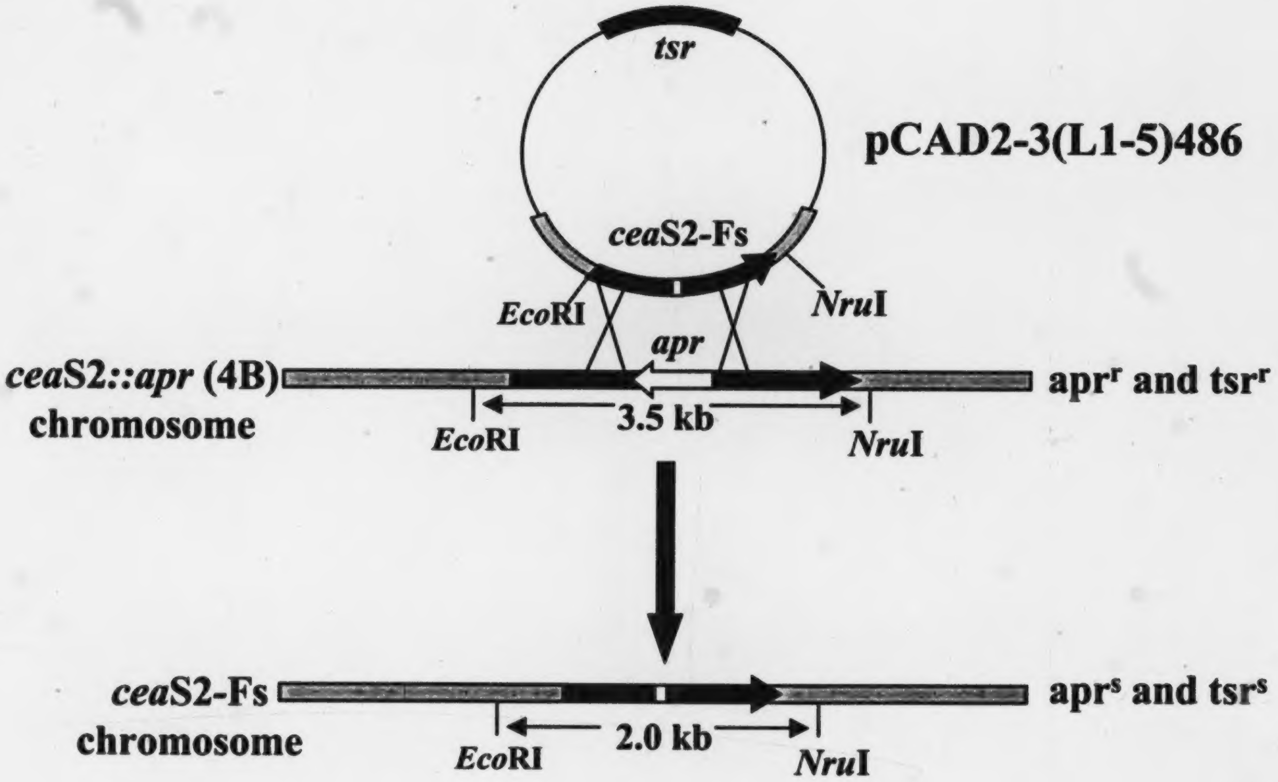

GCGGCCGC
CGCCGGCG
pCAD2-3
NotI digestion
GC GGCCGC
CGCCGG CG
Filling in of recessive ends
GCGGCC GGCCGC
CGCCGG CCGGCG
Recircularize by ligation
GCGGCCGGCCGC
CGCCGGCCGGCG
+ 1 Frame Shift
pCAD2-3(L1-5)

Figure 3.1.2 Homologous recombination involving pCAD2-3(L1-5) and the *S. clavuligerus cea*S2::*apr* (4B) mutant chromosome. Protoplasts of a *S. clavuligerus* strain in which *cea*S2 was disrupted by the insertion an apramycin resistance gene cassette (*apr*), were transformed with pCAD2-3(L1-5), which carries *cea*S2 disrupted by a frame shift mutation (*cea*S2-Fs) (Figure 3.1.1). The *cea*S2 coding sequence is represented by the filled arrows and the surround chromosomal region is represented by the gray bars. The hollow arrow and the hollow bar within *cea*S2 represent *apr* and the frame shift mutation, respectively, and the filled bar on pCAD2-3(L1-5) represents the thiostrepton resistance gene (*tsr*).


tsr
pCAD2-3(L1-5)486
*cea*S2-Fs
*Eco*RI
*Nru*I
apr
*cea*S2::*apr* (4B) chromosome
apr^r and tsr^r
*Eco*RI
3.5 kb
*Nru*I
*cea*S2-Fs chromosome
apr^s and tsr^s
*Eco*RI
2.0 kb
*Nru*I

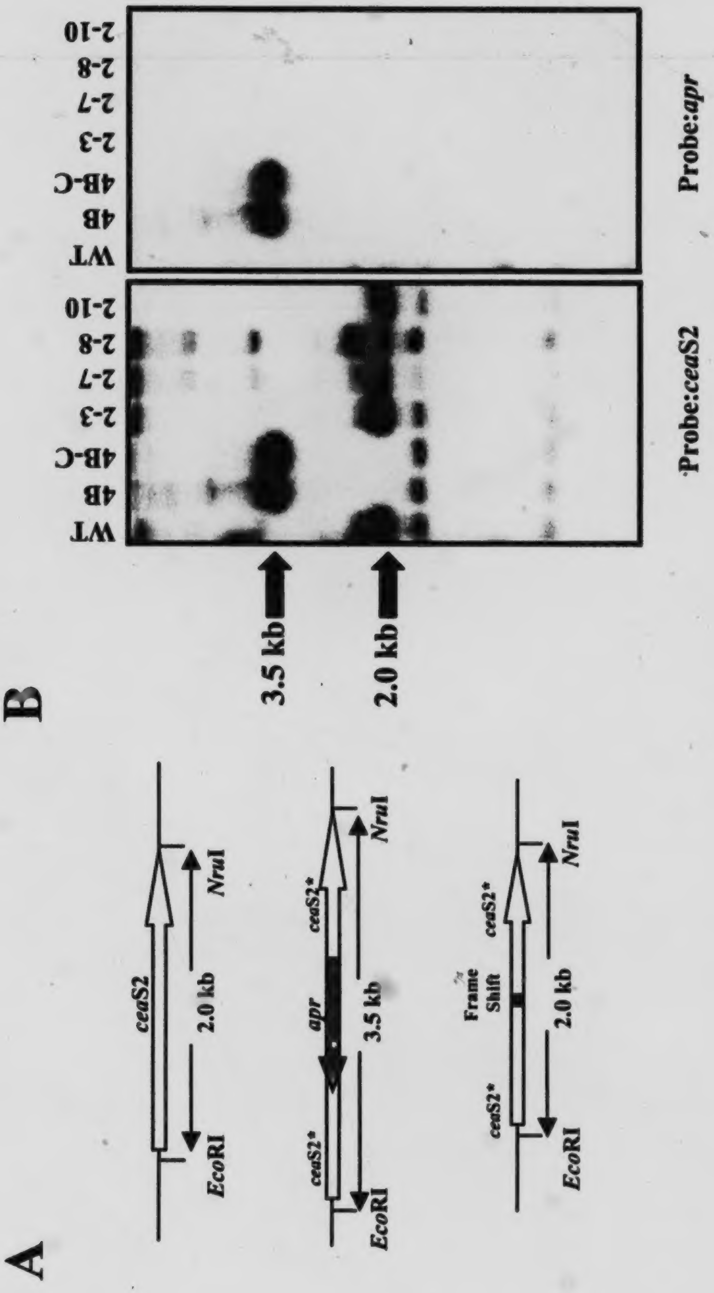
was passed through *S. clavuligerus* MEL and was then transformed into the *S. clavuligerus ceaS2::apr* (4B) mutant (Table 2.2). The *ceaS2::apr* (4B) mutant has a disruption in the *ceaS2* gene resulting from the insertion of an apramycin resistance gene cassette, in the orientation opposite to *ceaS2* transcription (Figure 3.1.2)(Jensen *et al.*, 2000). Apramycin and thiostrepton resistant transformants (apr^r and tsr^r respectively) were selected and allowed to sporulate twice on non-selective medium to isolate apr^s and tsr^s mutants, resulting from homologous recombination leading to gene replacement (Figure 3.1.2).

To confirm the replacement of the chromosomal *apr*-disrupted copy of *ceaS2* by the plasmid-encoded *ceaS2*-Fs copy, genomic DNA from the new *ceaS2*-Fs mutants, the parental *ceaS2::apr* mutant and the wild type strain was analyzed by Southern hybridization after digestion with *Eco*RI and *Nru*I. When an 855 bp *Eco*RI-*Not*I fragment from pCAD2-3, which contained the 5′ region of *ceaS2* and some upstream sequence, was used as a *ceaS2*-specific probe, a 2.0 kb fragment hybridized to the probe in the wild type and the new *ceaS2*-Fs mutant sample lanes (Figure 3.1.3). In contrast, the same probe hybridized to a 3.5 kb fragment in lanes containing DNA from the parental *ceaS2::apr* mutants (Figure 3.1.3). When the same blot was stripped and re-probed using a 1.45 kb *Nco*I fragment containing *apr* from pUC120Apr as the probe, no hybridizing bands were seen in lanes containing DNA from the wild type and the *ceaS2*-Fs mutants, whereas a 3.5 kb fragment hybridized to the probe in lanes containing DNA from the *ceaS2::apr* mutant (Figure 3.1.3). These results were consistent with the replacement of *ceaS2::apr* by *ceaS2*-Fs on the *S. clavuligerus* chromosome.

3.1.1.2 Phenotype of the *ceaS2*-Fs mutant

Four *ceaS2*-Fs mutants were fermented in SA and soy media in single shake flask cultures, along with wild type and parental *ceaS2::apr* strains. Supernatants from 72 hour and 96 hour cultures were analyzed by HPLC and bioassays for clavulanic acid and 5*S* clavam metabolite production. After 72 hours of growth in soy fermentation medium, clavulanic acid was detected in culture supernatants from the wild type strain, but only one *ceaS2*-Fs mutant (2-10) produced clavulanic acid and that too, only in trace amounts.

Figure 3.1.3 Southern analysis of *cea*S2-Fs mutants. (A) Diagrammatic representation of the *cea*S2 region of the *S. clavuligerus* chromosome in the wild type, the *cea*S2*::apr* (4B) mutant and the *cea*S2-Fs mutants. The gray arrow represents the *apr* disruption cassette, and the open arrow represents *cea*S2 with the direction of transcription represented by the arrowheads. The solid bar represents the frame shift mutation, the fine lines represent the rest of the *S. clavuligerus* chromosome and the disrupted *cea*S2 coding sequence is represented by (*). (B) Southern analysis of *Eco*RI and *Nru*I digested genomic DNA from *S. clavuligerus* wild type and *cea*S2 mutant strains. DNA from wild type, *cea*S2*::apr* mutants 4B and 4B-C, and from *cea*S2-Fs mutants 2-3, 2-7, 2-8 and 2-10 was probed using a *cea*S2-specific probe and an *apr*-specific probe.


A
ceaS2
2.0 kb
EcoRI
NruI
ceaS2*
apr
ceaS2*
3.5 kb
EcoRI
NruI
ceaS2*
Frame Shift
ceaS2*
2.0 kb
EcoRI
NruI
B
3.5 kb
2.0 kb
WT
4B
4B-C
2-3
2-7
2-8
2-10
Probe:ceaS2
WT
4B
4B-C
2-3
2-7
2-8
2-10
Probe:apr

HPLC analysis of soy culture supernatants after 96 hours of growth showed that two more of the *cea*S2-Fs mutants now produced some clavulanic acid, but levels were still very low at ~11% as compared to the wild type strain (Table 3.1). Varying amounts of 5*S* clavam metabolites were also detected in these culture supernatants (Table 3.2), but production of all of the detectable clavam metabolites was depressed relative to the wild type control (Figure 3.1.4). As demonstrated in previous studies (Jensen *et al.*, 2000), both the *cea*S2*::apr* and *cea*S2-Fs mutants were completely blocked in clavulanic acid and clavam metabolite biosynthesis in SA medium (Figure 3.1.4). In addition, bioassays indicated that all of the *cea*S2 mutants produced cephamycin C at similar levels to that produced by the wild type strain.

The results obtained by fermenting the *cea*S2-Fs mutants were inconsistent, as all of the mutants displayed a variable phenotype in terms of clavulanic acid and 5*S* clavam production levels. When fermentation studies were repeated using the same mutants, some of them still produced the respective metabolites whereas the others did not as determined by HPLC analysis and bioassays of soy culture supernatants. This unpredictability persisted and depended on the specific fermentation experiment, although all the *cea*S2-Fs mutants (except for 2-8) did produce the respective metabolites at some point, as detected by HPLC analysis.

To obtain better sensitivity, LC-MS analysis of culture supernatants after imidazole derivation was also carried out. LC-MS analysis of soy culture supernatants showed the presence of clavam-2-carboxylate, alanylclavam, 2-hydroxymethyl clavam and clavulanic acid fragmented ions, in supernatants from the wild type strain (Figure 3.1.5). The detailed mass spectra of the respective metabolites were also analyzed to confirm their identities (Figure 3.1.6). When the *cea*S2-Fs mutants were analyzed, the greater sensitivity of the LC-MS procedure meant that clavulanic acid was now detected in all mutants including *cea*S2-Fs (2-8) (Figure 3.1.7), which had previously been negative by HPLC alone (Tables 3.1 and 3.2). This showed that the *cea*S2-Fs (2-8) mutant probably produces clavulanic acid at very low levels, indicating that all the *cea*S2-Fs mutants retain the ability to produce clavulanic acid to some degree.

A representative *cea*S2-Fs mutant (2-3) and the wild type strain were also fermented using GSPG fermentation medium as it has been reported by others that *cea*S2

Table 3.1: Clavulanic acid produced by *S. clavuligerus cea*S2*::apr* and *cea*S2-Fs mutants after 96 hours of growth on soy medium[a] as determined by HPLC[b] analysis of culture supernatants.

Strain	Clavulanic acid (μg/ml)	Clavulanic acid produced (% of wild type)[c]
Wild type	138.5	100
*cea*S2*::apr* (4B)	ND[d]	0
*cea*S2-Fs (2-3)	14.4	10.4
*cea*S2-Fs (2-7)	7.6	5.4
*cea*S2-Fs (2-8)	ND[d]	0
*cea*S2-Fs (2-10)	15.7	11.3

[a] Combined results from four independent fermentations are shown
[b] Results for the wild type and some *cea*S2-Fs strains were also verified by LC-MS
[c] Clavulanic acid produced by the wild type strain was assigned a value of 100%
[d] None detected

Table 3.2: Clavaminic acid, clavam-2-carboxylate and 2-hydroxymethyl clavam produced by *S. clavuligerus cea*S2::*apr* and *cea*S2-Fs mutants after 96 hours of growth on soy medium[a] as determined by HPLC[b] analysis of culture supernatants.

Strain	Clavaminic acid (µg/ml)/(% of wild type)[c]	Clavam-2-carboxylate (µg/ml)/(% of wild type)[c]	2-hydroxymethyl clavam (µg/ml)/(% of wild type)[c]
Wild type	32.3/100	83.5/100	75.2/100
*cea*S2::*apr* (4B)	ND[d]/0	ND[d]/0	ND[d]/0
*cea*S2-Fs (2-3)	9.3/28.8	27.0/32.3	26.9/35.8
*cea*S2-Fs (2-7)	0.07/0.2	1.2/1.4	2.5/3.3
*cea*S2-Fs (2-8)	ND[d]/0	ND[d]/0	ND[d]/0
*cea*S2-Fs (2-10)	1.2/3.7	5.6/6.7	5.5/7.3

[a] Combined results from four independent fermentations are shown
[b] Results for the wild type and some *cea*S2-Fs strains were also verified by LC-MS
[c] Production of the respective metabolites by the wild type strain was assigned a value of 100%
[d] None detected

Figure 3.1.4 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, *cea*S2*::apr* (4B) and *cea*S2-Fs mutant strains. Soy (A) and SA (B) culture supernatants from the respective strains were analyzed by HPLC. Of the four *cea*S2-Fs mutants fermented, chromatograms for only (2-10), which produced the maximum amounts of the respective metabolites are shown. The wild type samples were diluted (1/5) before analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.

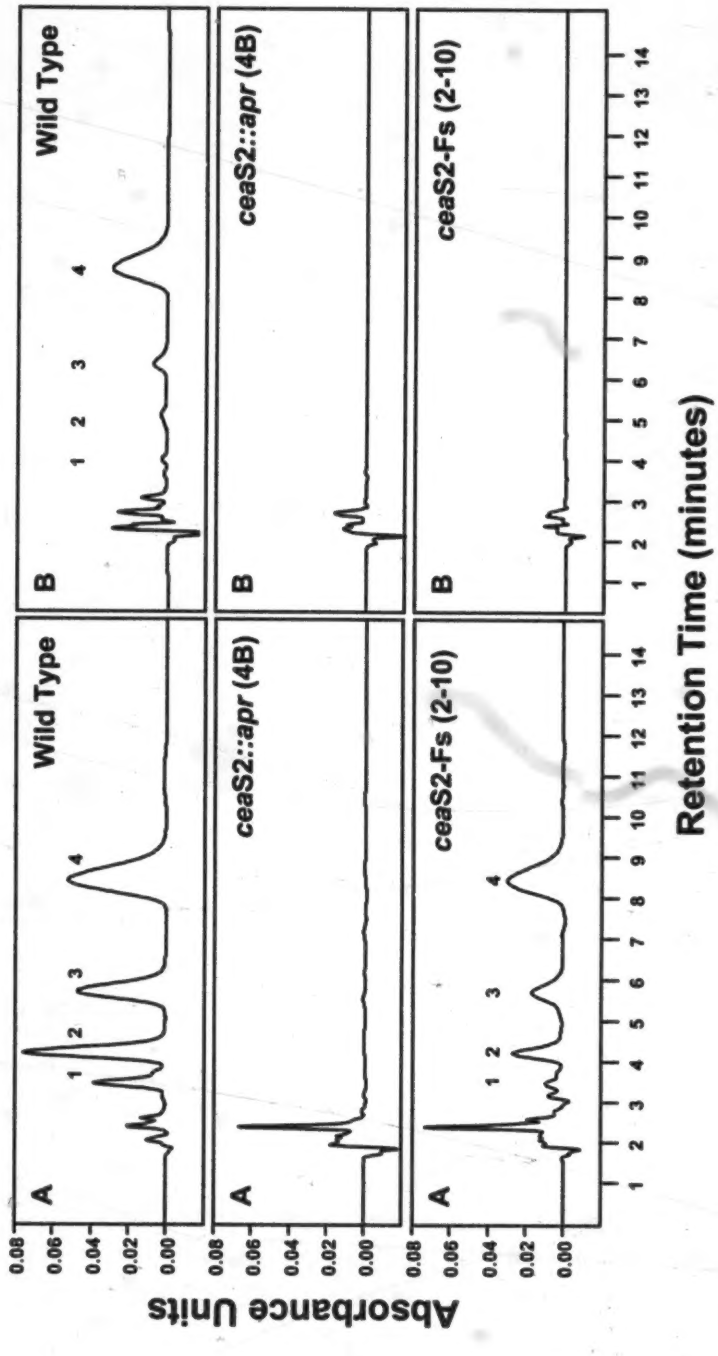

A
Wild Type
1
2
3
4
B
Wild Type
1
2
3
4
A
ceaS2::*apr* (4B)
B
ceaS2::*apr* (4B)
A
ceaS2-Fs (2-10)
1
2
3
4
B
ceaS2-Fs (2-10)
0.08
0.06
0.04
0.02
0.00
Absorbance Units
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Retention Time (minutes)

Figure 3.1.5 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type strain as determined by LC-MS analysis. Ninety six-hour soy culture supernatant from the wild type strain was analyzed by LC-MS. The chromatograms shown are as follows: (A) Absorbance at 311 nm due to clavam compounds derivatized with imidazole, (B) Peak due to derivatized alanylclavam fragmented ion, (C) Peak due to derivatized clavam-2-carboxylate fragmented ion, (D) Peak due to derivatized clavulanic acid fragmented ion and (E) Peak due to derivatized 2-hydroxymethyl clavam fragmented ion. The specific peaks observed due to the presence of the respective metabolites are indicated by filled arrows. For identification, the mass/charge (m/z) values for each metabolite was determined as described in Figures 3.1.6. The x-axis indicates the retention time and the y-axis indicates the relative abundance of the respective peaks.

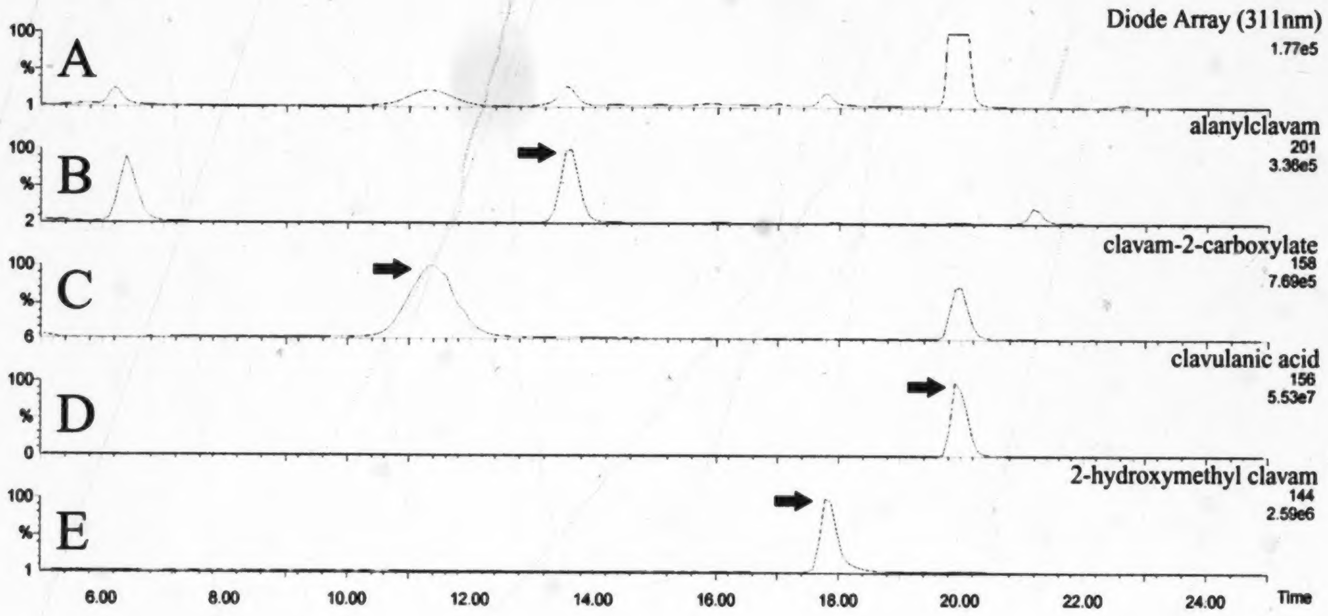
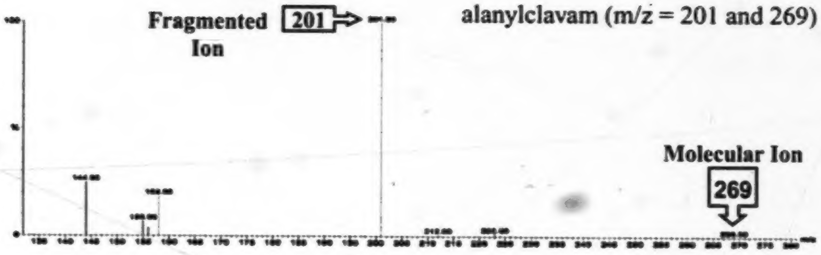
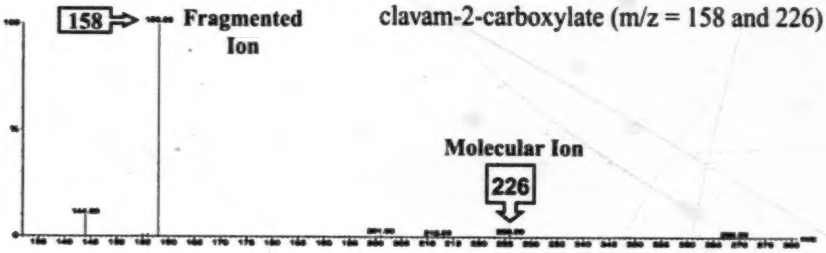

Diode Array (311nm)
1.77e5
A
alanylclavam
201
3.38e5
B
clavam-2-carboxylate
158
7.69e5
C
clavulanic acid
156
5.53e7
D
2-hydroxymethyl clavam
144
2.59e6
E
100
%
1
2
6
0
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time

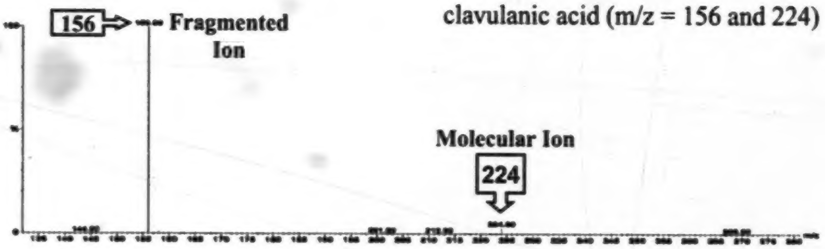
Figure 3.1.6 Mass spectra of different clavam compounds analyzed by LC-MS. The characteristic mass spectra of imidazole derivatized clavam-2-carboxylate, alanylclavam, 2-hydroxymethyl clavam and clavulanic acid, indicating the molecular and fragmented ions are shown. The corresponding ions were used to identify and verify the identities of the respective metabolites being subjected to analysis by LC-MS. The x-axis indicates the (m/z) ratio and the y-axis indicates the relative abundance of the respective ions.



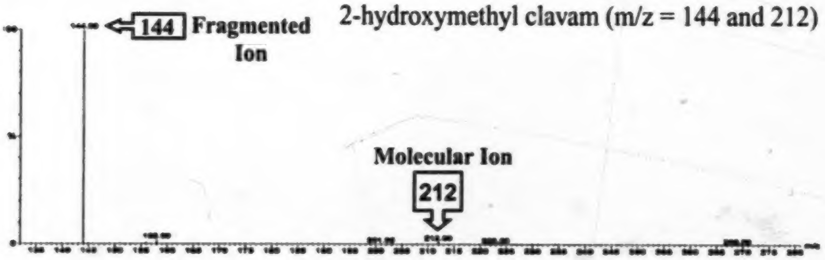
alanylclavam (m/z = 201 and 269)



clavam-2-carboxylate (m/z = 158 and 226)

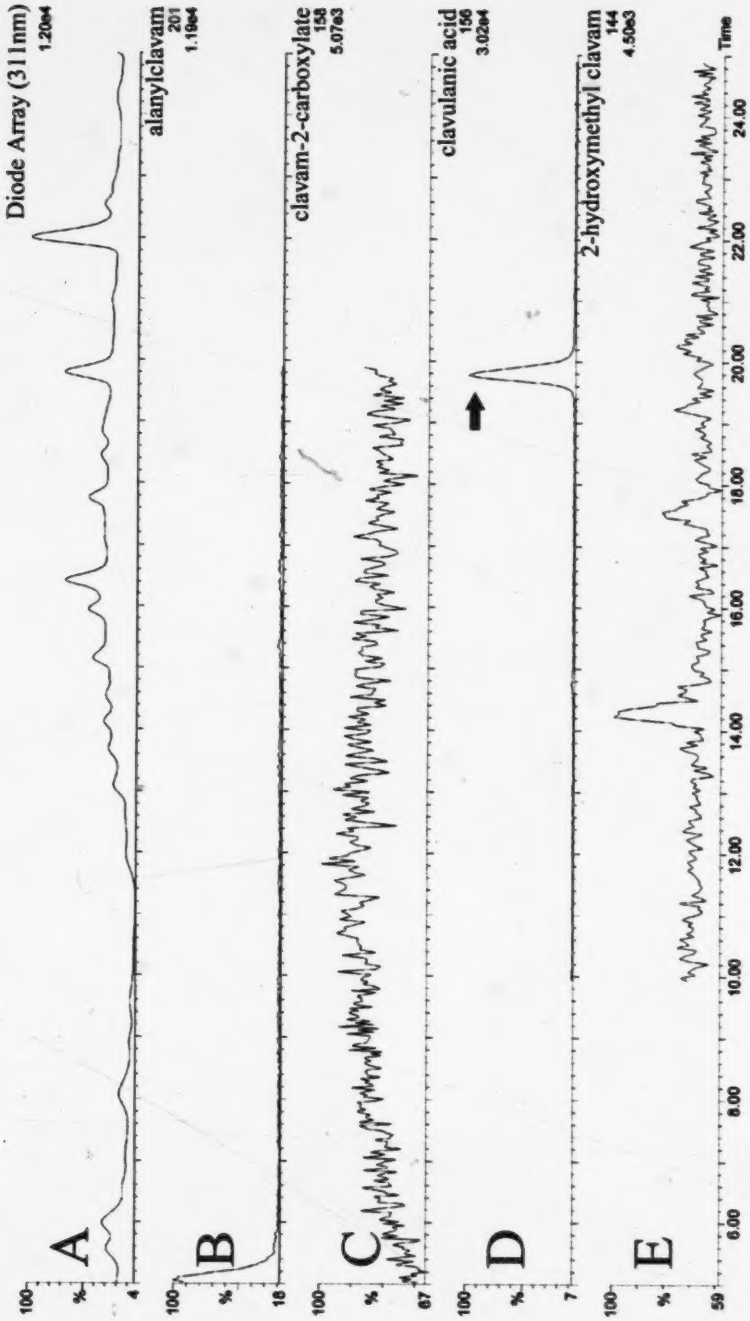


clavulanic acid (m/z = 156 and 224)



2-hydroxymethyl clavam (m/z = 144 and 212)

Figure 3.1.7 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus cea*S2-Fs (2-8) mutant as determined by LC-MS. Culture supernatant from 96-hour soy culture was analyzed by LC-MS. The chromatograms shown are as follows: (A) Absorbance at 311 nm due to clavam compounds derivatized with imidazole, (B) Profile due to derivatized alanylclavam fragmented ion, (C) Profile due to derivatized clavam-2-carboxylate fragmented ion, (D) Peak due to derivatized clavulanic acid fragmented ion, and (E) Profile due to derivatized 2-hydroxymethyl clavam fragmented ion. The peaks corresponding to the respective metabolites are indicated by filled arrows. For identification, the mass/charge (m/z) values for each metabolite were analyzed in detail as described earlier (Figures 3.1.6). The x-axis indicates the retention time and the y-axis indicates the relative abundance of the respective peaks.


Diode Array (311nm)
1.20e4
alanylclavam
201
1.19e4
clavam-2-carboxylate
158
5.07e3
clavulanic acid
156
3.02e4
2-hydroxymethyl clavam
144
4.50e3
A
B
C
D
E
100
%
4
100
%
18
100
%
67
100
%
7
100
%
59
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time

mutants only produce clavulanic acid and the 5*S* clavams when grown in this medium (Perez-Redondo *et al.*, 1999). Culture supernatants were withdrawn from wild type and *cea*S2-Fs (2-3) GSPG grown cultures after 24, 48, 72, 96 and 120 hours of growth, and were analyzed by HPLC. The wild type strain produced the maximum levels of the respective metabolites after 48 hours of growth in GSPG medium (Table 3.3). On the other hand, the *cea*S2-Fs (2-3) mutant did not produce any detectable levels of clavulanic acid or the 5*S* clavams under the same conditions, at any of the time points analyzed (Table 3.3).

It has also been suggested that the production of clavulanic acid by the *cea*S2 mutant is dependent on the presence of glycerol in the growth medium (Perez-Redondo *et al.*, 1999). Therefore, the wild type strain and the *cea*S2-Fs (2-3) mutant were fermented in SA and soy media supplemented with 15 g/l glycerol. Cultures were fermented for 72, 96 and 120 hours and were analyzed by bioassay for clavulanic acid and alanylclavam biosynthesis. Both of the respective metabolites were only detected in culture supernatants from the wild type strain and not in supernatants from the *cea*S2-Fs cultures.

Because the *cea*S2*::apr* and the *cea*S2-Fs mutants were so unpredictable with respect to their ability to produce clavulanic acid and the 5*S* clavams, and even when they did show some low level production, it was hard to conclude reliably that a second functional copy a *cea*S gene must be present. In a final attempt to establish conditions under which the *cea*S2-Fs mutants are consistently able to produce clavulanic acid and the 5*S* clavams, one *cea*S2-Fs (2-3) mutant and the wild type strain were fermented in SA and soy media supplemented with dehydroacetic acid (DHA). DHA has been shown to increase the levels of clavulanic acid and the 5*S* clavams produced by *S. clavuligerus* under some growth conditions by a still unknown mechanism (S.E. Jensen, Personal communication). Therefore, if the *cea*S2 mutants had any genuine ability to produce the clavam metabolites, it should be enhanced in the presence of DHA. Cultures (SA and soy) were propagated and supernatants were analyzed after 72, 96 and 120 hours of growth as described earlier. In cultures supplemented with DHA, it was found that the *cea*S2-Fs mutant produced clavulanic acid in SA medium and clavam-2-carboxylate, 2-hydroxymethyl clavam and clavulanic acid in soy medium, but still in very low amounts relative to the wild type strain (Table 3.4).

Table 3.3: Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type and *cea*S2-Fs, mutant strains grown in glycerol containing medium.

Strain[a]	GSPG[b]		SA-G[c]		Soy-G[c]	
	5*S* clavams[d]	Clavulanic acid	5*S* clavams[d]	Clavulanic acid	5*S* clavams[d]	Clavulanic acid
Wild type[e]	+	+	-	+	+	+
*cea*S2-Fs (2-3)[e]	-	-	-	-	+/-	-

[a] Results were verified by fermenting two separate *cea*S2-Fs mutants

[b] Supernatants from cultures grown on GSPG medium for 48 hours were analyzed by HPLC

[c] Supernatants from cultures grown in SA or soy medium for 96 hours, supplemented with 15 mg/l glycerol (SA-G and Soy-G, respectively) were analyzed by bioassays

[d] The clavams are shown as a family of compounds as the bioassays only detected alanylclavam production

[e] Qualitative results are shown: +, the respective metabolites were detected; -, the respective metabolites were not detected and +/-, non-significant zones of clearing were obtained therefore no clear conclusion could be drawn

Table 3.4: Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type and *cea*S2-Fs mutant strains as determined by HPLC analysis of culture supernatants, after 96 hours of growth in SA and soy media supplemented with dehydroacetic acid (DHA).

Strain[a]	**Clavaminic acid**		**Clavam-2-carboxylate**		**2-hydroxymethyl clavam**		**Clavulanic acid**	
				(µg/ml)				
	SA-D[b]	**Soy-D**[b]	**SA-D**[b]	**Soy-D**[b]	**SA-D**[b]	**Soy-D**[b]	**SA-D**[b]	**Soy-D**[b]
Wild type	0.17	18.7	1.52	50.42	1.86	92.25	1.73	117.45
*cea*S2-Fs (2-3)	ND[c]	ND[c]	ND[c]	2.41	ND[c]	2.68	0.02	5.86

[a] Results were verified by fermenting two separate *cea*S2-Fs mutants

[b] SA or soy medium supplemented with 0.5 mg/ml DHA (SA-D and Soy-D, respectively)

[c] None detected

3.1.1.3 Complementation of the *cea*S2-Fs mutant

To determine if the variable phenotype of the *cea*S2-Fs mutants in terms of clavulanic acid and 5*S* clavam metabolite biosynthesis was only due to the mutation introduced into *cea*S2, the *cea*S2-Fs (2-8) mutant strain was complemented using a plasmid encoding the entire wild type *cea*S2 gene. The *cea*S2-Fs (2-8) mutant was selected for complementation, as it never produced clavulanic acid at high enough levels to be detected by HPLC. The plasmid pSET152-5.3, which contains the *cea*S2 gene along with its upstream region as a 5.3 kb *Bam*HI/*Xba*I fragment, ligated into the corresponding sites of the *Streptomyces* integrative vector pSET152 (Table 2.3), was introduced into the *cea*S2-Fs (2-8) mutant by conjugation. As a control, the plasmid vector pSET152 without any insert was also introduced into the *cea*S2-Fs (2-8) mutant. Strains containing the respective plasmids were isolated based on the apr^r marker present on the vector pSET152.

One control [*cea*S2-Fs (2-8) containing pSET152 only], two complemented strains (7 and 8) along with the parental *cea*S2-Fs (2-8) mutant and the wild type strain were fermented in SA and soy media. After 72 and 96 hours of growth on soy medium, the complemented mutant (#7) produced 2.9% and 6.1% clavulanic acid as compared to the wild type strain but no 5*S* clavam production was seen. The second complemented mutant fermented (#8), and the pSET152-containing control strain did not produce any detectable levels of the respective metabolites. In SA medium, only the wild type strain produced detectable levels of clavulanic acid, whereas all the other strains tested were blocked in clavulanic acid biosynthesis, as determined by HPLC.

3.1.2. Preparation and analysis of the *cea*S2 deletion mutant

Although production levels were unpredictable, *cea*S2-Fs mutants generally produced higher levels of clavulanic acid and the 5*S* clavams then the *cea*S2*::apr* mutants as determined by HPLC analysis of culture supernatants, indicating that some polar effects may have been relieved (Figure 3.1.4). Even then, some ambiguities arose as the levels of clavulanic acid and the 5*S* clavams produced by the *cea*S2-Fs mutants showed a great deal of variation, and in some cases the mutants did not produce the respective metabolites at all (Section 3.1.1.2). Since the *cea*S2-Fs mutant has a frame shift mutation in *cea*S2, this may give rise to a truncated protein, 463 amino acids in length, in comparison to the 573 amino acid wild type CeaS2 (Figure 3.1.8). Only the first 227 amino acids are likely to be the same in the wild type and the truncated forms of CeaS2 as a change in the reading frame, 679 bp from the translational start codon in *cea*S2-Fs, leads to a garbling of the amino acid sequence past the frame shift mutation (Figure 3.1.8). Like the *cea*S2-Fs mutants, the *cea*S2*::apr* mutants from which they were derived could also potentially give rise to truncated CeaS2 proteins, since the *cea*S2-Fs mutant was prepared using the same *Not*I restriction site that was used to prepare the original *cea*S2*::apr* mutant (Jensen *et al.*, 2000). Therefore, it was hypothesized that the truncated form of CeaS2 may somehow interfere with the biosynthesis of clavulanic acid and the 5*S* clavams by interacting with other components present in the pathway, leading to some form of negative dominance (L.W. Frost, Personal communication). To address this question, a new *cea*S2 mutant was prepared in which the entire gene was deleted and replaced by an apr^r gene cassette (Δ*cea*S2*::apr*), thereby eliminating the possibility of the expression of any truncated CeaS2 protein.

3.1.2.1 Preparation of the Δ*cea*S2*::apr* mutant

The *S. clavuligerus* Δ*cea*S2*::apr* mutant strain was prepared using the REDIRECT© PCR targeting system developed initially for the preparation of gene mutations in *S. coelicolor* (Gust *et al.*, 2003). This system utilizes the λ RED (*gam*, *bet* and *exo*) functions (Datsenko and Wanner, 2000) to promote homologous recombination between a linear PCR product and circular cosmid DNA, in order to disrupt cosmid-encoded genes. The mutagenized cosmid then functions as the disruption construct for isolating *Streptomyces* gene mutants. The primers ceaS2-FOR and ceaS2-REV were used to PCR-amplify the REDIRECT© disruption cassette from the template plasmid pIJ773. The resulting PCR product comprised the *aac(3)IV* gene, conferring apr^r, and a RK2/RP4 origin of transfer (*oriT*), both of which were flanked by FRT (FLP Recombination Target) sites (Section 3.1.3.1) and DNA sequence homologous to regions immediately upstream and downstream of *cea*S2 (Figure 3.1.9). The cosmid 12B8, which encompasses the region of the *S. clavuligerus* chromosome extending from ~3.0 kb upstream to ~40 kb downstream of *cea*S2, was transformed into the *E. coli*

Figure 3.1.8 Putative truncated form of CeaS2, arising due to the *cea*S2-Fs mutation. (A) The hollow arrows represent the wild type *cea*S2, the *cea*S2-Fs mutant and the *cas*2 genes, with the arrowheads representing the direction of transcription. The tips of the arrowheads indicate the translational stop codons present in the wild type copies of the respective genes. The black bar in *cea*S2-Fs represents the frame shift mutation (FS) in *cea*S2 and the gray bar represents the new translational stop codon (SC) introduced due to frame shifting. The 5.3 kb polycistronic transcript encoding the early genes from the clavulanic gene cluster is shown, and the hatched boxes represent the wild type (full length) and truncated forms of the CeaS2 protein. The sizes in bp of the ORFs corresponding to wild type *cea*S2 and *cea*S2-Fs are also indicated and the additional ORFs between *cea*S2 and *cas*2 are not shown for the sake of simplicity. (B) Comparison between the putative proteins encoded by *cea*S2 and *cea*S2-Fs. Conserved residues are shown on black background, similar residues are shown on gray background and different residues are shown on white background. The site of the introduced frame shift mutation in *cea*S2-Fs, resulting in the different amino acid sequences of CeaS2 and CeaS2-Fs is indicated by the filled inverted triangle.

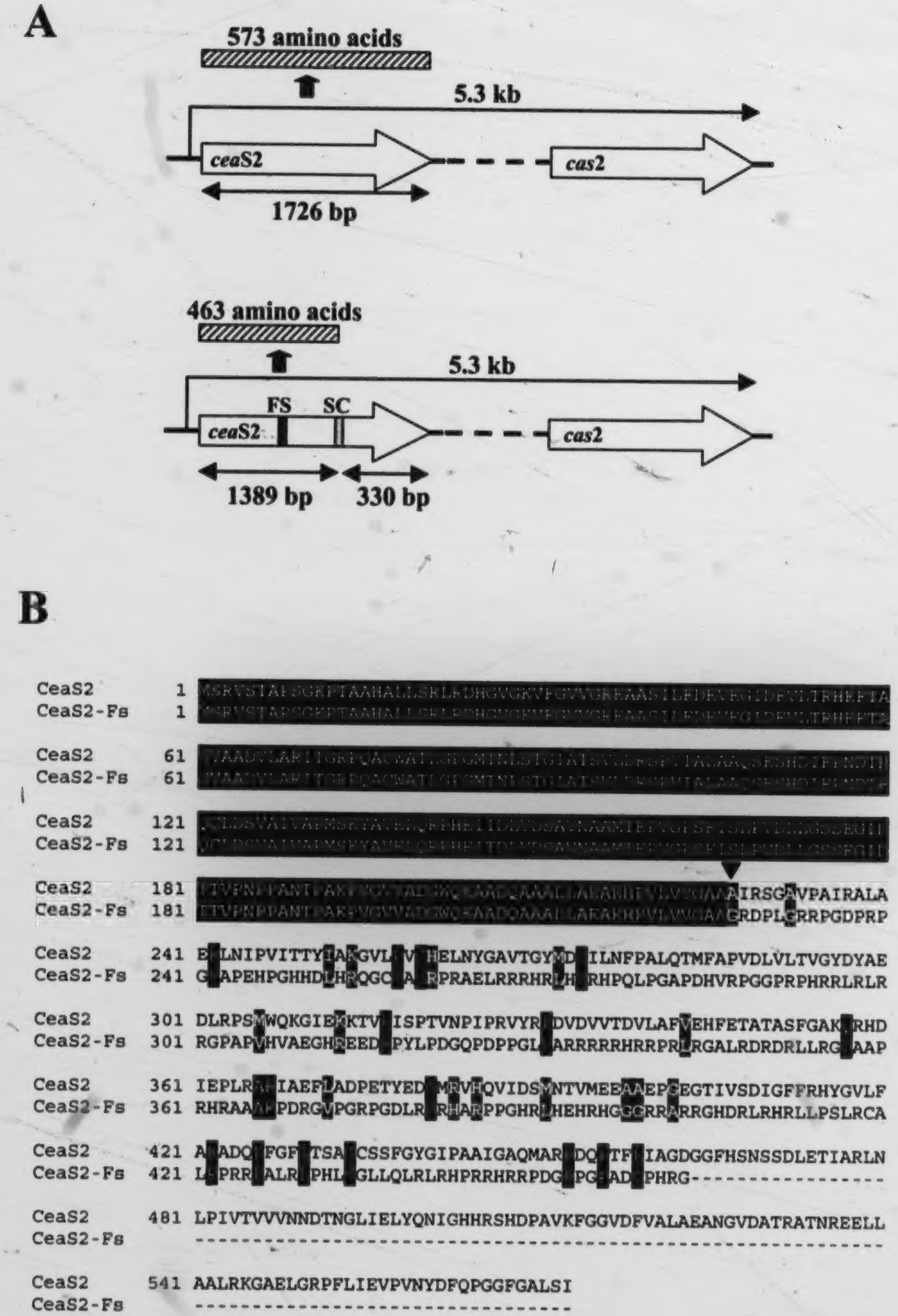
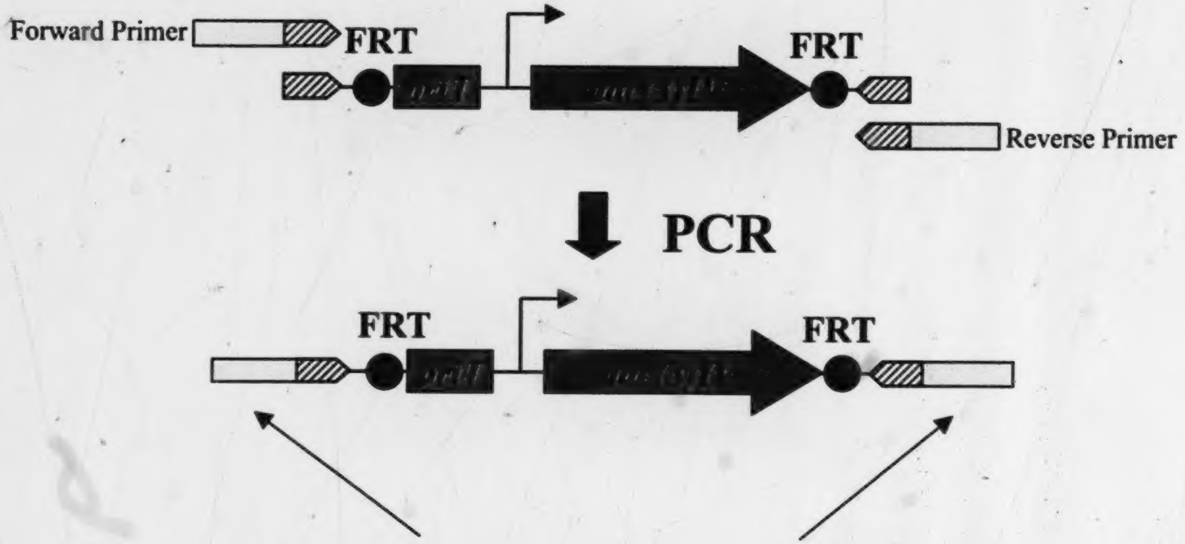

A
573 amino acids
5.3 kb
ceaS2
cas2
1726 bp
463 amino acids
5.3 kb
FS
SC
ceaS2
cas2
1389 bp
330 bp
B
CeaS2 1
CeaS2-Fs 1
CeaS2 61
CeaS2-Fs 61
CeaS2 121
CeaS2-Fs 121
CeaS2 181 AIRSGAVPAIRALA
CeaS2-Fs 181 GRDPLGRRPGDPRP
CeaS2 241 ELNIPVITTYLAKGVLAVNHELNYGAVTGYMDAILNFPALQTMFAPVDLVLTVGYDYAE
CeaS2-Fs 241 GAPEHPGHHDLHRQGCAARPRAELRRRHRLHARHPQLPGAPDHVRPGGPRPHRRLRLR
CeaS2 301 DLRPSWWQKGIEKKTVAISPTVNPIPRVYRADVDVVTDVLAFVEHFETATASFGAKARHD
CeaS2-Fs 301 RGPAPVVHVAEGHREEDAPYLPDGQPDPPGLAARRRRRHRRPRLRGALRDRDRLLRGAAAP
CeaS2 361 IEPLRAAIAEFLADPETYEDAMRVHQVIDSMNTVMEEAAEPGEGTIVSDIGFFRHYGVLF
CeaS2-Fs 361 RHRAAAAPDRGVPGRPGDLRARHARPPGHRAHEHRHGGCRRARRGHDRLRHRLLPSLRCA
CeaS2 421 AAADQAFGFATSALCSSFGYGIPAAIGAQMARADQATFAIAGDGGFHSNSSDLETIARLN
CeaS2-Fs 421 LAPRRAALRAPHLAGLLQLRLRHPRRHRRPDGAPGAADAPHRG----------------
CeaS2 481 LPIVTVVVNNDTNGLIELYQNIGHHRSHDPAVKFGGVDFVALAEANGVDATRATNREELL
CeaS2-Fs --
CeaS2 541 AALRKGAELGRPFLIEVPVNYDFQPGGFGALSI
CeaS2-Fs ---------------------------------

Figure 3.1.9 Preparation of the gene disruption cassette using REDIRECT© technology. The disruption cassette from the plasmid pIJ773 was isolated as a 1384 bp *Eco*RI/*Hin*dIII fragment and used as the template in a PCR reaction. The forward and reverse primers contained 39 bp nucleotide tails (shown as gray boxes) that were homologous to sequences immediately upstream and downstream of the gene to be disrupted (*cea*S2), respectively. The hatched arrows represent the priming sequences based on the plasmid pIJ773. The FLP recombination target sites (FRT) are shown as filled circles. The bent arrow indicates the promoter upstream of the apramycin resistance gene [*acc(3)IV*] and the arrowhead represents the direction of transcription. The broad host range RK2/RP4 origin of transfer (*oriT*) included in the disruption cassette is also shown.


Forward Primer
FRT
FRT
Reverse Primer
PCR
FRT
FRT

DNA sequence homologous to regions immediately upstream and downstream of *cea*S2

BW25113/pIJ780 strain (Table 2.1), and then the PCR-generated REDIRECT© disruption cassette was also introduced by electroporation. The plasmid pIJ780 in *E. coli* BW25113 encodes the λ RED proteins, which are expressed from an arabinose inducible promoter. Homologous recombination between the linear PCR product and the cosmid 12B8 resulted in the entire *cea*S2 gene being deleted from the cosmid and replaced by the disruption cassette (Δ*cea*S2::*apr*), to give the mutant cosmid 12B8-AP (Figure 3.1.10). The cosmid 12B8-AP was then introduced into wild type *S. clavuligerus* by conjugation as it now contained an origin of transfer, which was a part of the disruption cassette, and exconjugants were selected based on apr^r and kanamycin sensitivity (kan^s) to isolate Δ*cea*S2::*apr* mutant strains (Figure 3.1.11).

Southern analysis was used to confirm that the chromosomal wild type copy of *cea*S2 had been replaced by the antibiotic resistance gene cassette. Genomic DNA from the wild type strain and two Δ*cea*S2::*apr* mutants (#2 and #3) was digested with *Nco*I, and was used in the analysis (Figure 3.1.12A). When membranes with bound DNA were hybridized to a *cea*S2-specific probe consisting of the 855 bp *Eco*RI-*Not*I fragment from pCAD2-3, a 2.87 kb *Nco*I fragment hybridized to the probe in lanes containing DNA from the wild type strain, whereas the corresponding band was not observed in lanes containing DNA from the Δ*cea*S2::*apr* mutants (Figure 3.1.12B). When the same blot was stripped and hybridized to the 1.38 kb *Eco*RI-*Hin*dIII fragment from pIJ773, which encompassed the *aac(3)IV+oriT* disruption cassette, a 2.53 kb *Nco*I fragment hybridized to the probe in lanes containing DNA from the Δ*cea*S2::*apr* mutants, whereas no hybridization was observed in lanes containing DNA from the wild type strain (Figure 3.1.12B). These results were consistent with the deletion of the chromosomally encoded wild copy of *cea*S2, and its replacement with the *aac(3)IV+oriT* disruption cassette (Figure 3.1.12A).

3.1.2.2 Phenotype of the Δ*cea*S2::*apr* mutant

Two Δ*cea*S2::*apr* mutants that had been verified by Southern analysis were fermented in SA and soy media for 72 and 96 hours along with the wild type strain, and the culture supernatants were analyzed by HPLC. After 72 and 96 hours of growth, the wild type strain produced expected levels of clavulanic acid in SA and both clavulanic

Figure 3.1.10 Preparation of the mutagenized cosmid 12B8-AP containing the disrupted copy of *cea*S2. The REDIRECT© disruption cassette (Figure 3.1.9) was transformed into the *E. coli* BW25113/pIJ780 strain containing the cosmid 12B8. This strain also contains the plasmid pIJ780, which encodes the λ RED proteins, which are expressed from an arabinose inducible promoter. The λ RED proteins facilitate homologous recombination between the linear PCR product (disruption cassette) and the cosmid 12B8. This leads to the replacement of *cea*S2 by the *acc(3)IV+oriT* cassette in the mutant cosmid 12B8-AP. The *neo* gene conferring neomycin/kanamycin resistance (by cross resistance) and the ampicillin resistance gene (*amp*), which are present on the cosmid are also shown.

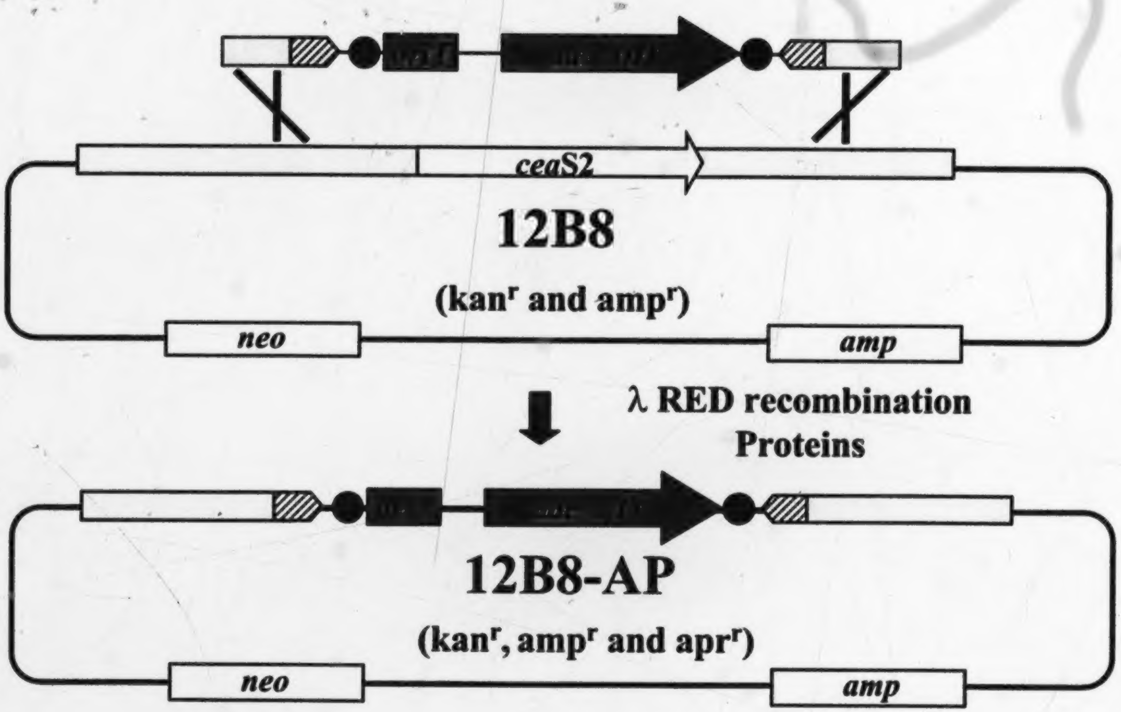
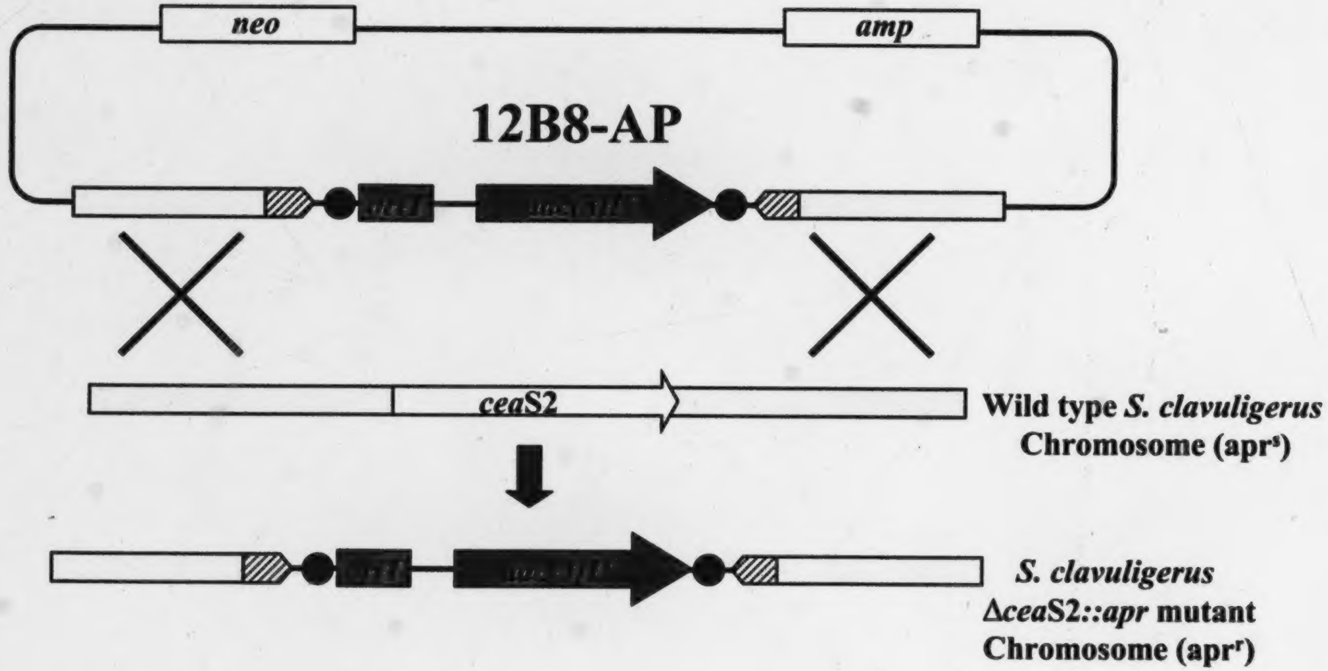
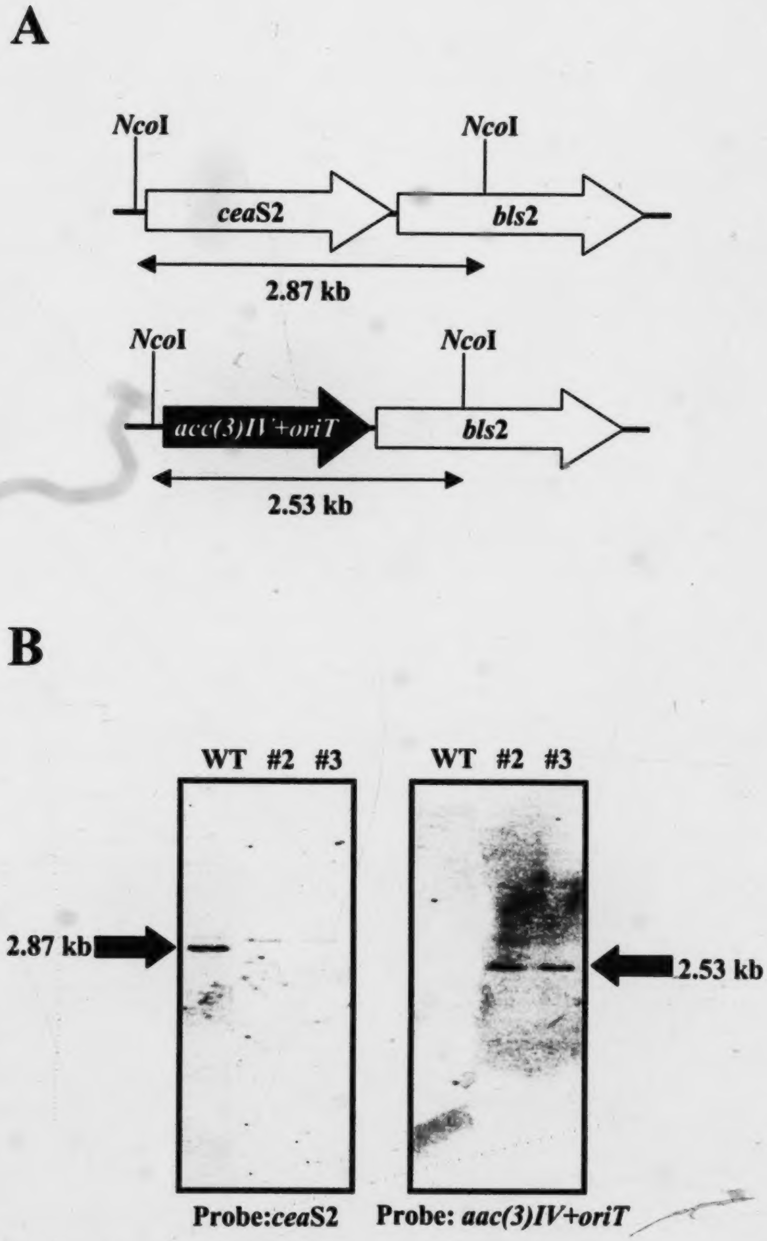

*cea*S2
12B8
(kan^r and amp^r)
neo
amp
λ RED recombination Proteins
12B8-AP
(kan^r, amp^r and apr^r)
neo
amp

Figure 3.1.11 Construction of the *S. clavuligerus* Δ*cea*S2::*apr* mutant using the mutagenized recombinant cosmid 12B8-AP. The cosmid 12B8-AP in which *cea*S2 has been deleted and replaced by the *acc(3)IV+oriT* cassette from the plasmid pIJ773 (Figure 3.1.10) was introduced into wild type *S. clavuligerus* by conjugation. Homologous recombination and double cross over events between the cosmid, which lacked a *Streptomyces* origin of replication, and the *S. clavuligerus* chromosome, led to the replacement of the chromosomal copy of *cea*S2 by the disruption cassette. This resulted in the isolation of apr^r and kan^s, Δ*cea*S2::*apr* mutants.



neo
amp
12B8-AP
*cea*S2
Wild type *S. clavuligerus* Chromosome (apr^s)
S. clavuligerus Δ*cea*S2::*apr* mutant Chromosome (apr^r)

Figure 3.1.12 Southern analysis of the *S. clavuligerus* Δ*cea*S2*::apr* mutants. (A) Diagrammatic representation of the *cea*S2 region of the *S. clavuligerus* chromosome in the wild type and the Δ*cea*S2*::apr* mutants. The open arrows represent *cea*S2 and *bls*2 with the direction of transcription represented by the arrowheads. The filled arrow represents the *acc(3)IV+oriT* cassette in the Δ*cea*S2*::apr* mutants and the fine lines represent the rest of the *S. clavuligerus* chromosome. (B) Southern analysis of *Nco*I digested genomic DNA from *S. clavuligerus* wild type (WT) and Δ*cea*S2*::apr* (#2 and #3) mutant strains. After processing, DNA from the respective strains was probed using a *cea*S2-specific and an *acc(3)IV+oriT*-specific probe. The additional restriction fragments hybridizing to the *cea*S2-specific probe arise due to the presence of extra DNA in the probe preparation, corresponding to the region surrounding *cea*S2.


A
NcoI
NcoI
ceaS2
bls2
2.87 kb
NcoI
NcoI
acc(3)IV+oriT
bls2
2.53 kb
B
WT #2 #3
WT #2 #3
2.87 kb
2.53 kb
Probe:ceaS2
Probe: aac(3)IV+oriT

acid and the 5*S* clavams in soy medium, respectively (Figure 3.1.13). Under the same conditions, both of the Δ*cea*S2*::apr* mutants fermented did not produce any of the respective metabolites in both media tested (Figure 3.1.13).

3.1.3 Preparation and analysis of the *cea*S2 in-frame deletion mutant

To complete the repertoire of different *cea*S2 mutants prepared in this study, a final mutant was prepared by eliminating the apr^r cassette from the Δ*cea*S2*::apr* mutant, to leave an unmarked in-frame deletion (Δ*cea*S2). This mutant should have the least potential to cause any polar effects on the expression of downstream genes, of all the *cea*S2 mutants analyzed so far.

3.1.3.1 Preparation of the *S. clavuligerus* Δ*cea*S2 mutant

The FRT (FLP Recombination Target) sites flanking the disruption cassette used for constructing the Δ*cea*S2*::apr* mutant by REDIRECT© PCR targeting, were used for preparing the *cea*S2 in-frame deletion mutant. A 5.89 kb *Kpn*I fragment from the cosmid 12B8-AP, encompassing the Δ*cea*S2*::apr* mutation, was subcloned into *Kpn*I-digested pUWL-KS to give pUWL-Δ*cea*S2*::apr* (Figure 3.1.14). The cosmid 12B8-AP and the plasmid pUWL-Δ*cea*S2*::apr* were then separately transformed into the DH5α(BT340) strain (Table 2.1), containing the plasmid BT340, which encodes the FLP recombinase expressed from a promoter induced at 42°C, and harbors a temperature-sensitive replicon (Datsenko and Wanner, 2000). The FLP recombinase is a site-specific recombinase which cleaves and re-ligates DNA at specific FRT sites (Cox, 1983). If two FRT sites are arranged in tandem, then FLP-mediated recombination at the two sites brings about the deletion of the DNA fragment flanked by the two sites (Cox, 1983). The disruption cassettes used in the REDIRECT© PCR targeting system are flanked by tandemly arranged FRT sites (Figure 3.1.9), and are designed as such so that FLP causes the excision of the antibiotic marker, leaving an 81 bp in-frame "scar" sequence (Figure 3.1.15)(Gust *et al.*, 2003). Therefore, when FLP expression was induced in the *E. coli* DH5α(BT340) strains containing either 12B8-AP or pUWL-Δ*cea*S2*::apr*, FLP mediated the excision of the *aac(3)IV+oriT* disruption cassette in both 12B8-AP and pUWL-

Figure 3.1.13 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* Δ*cea*S2::*apr* mutants as determined by HPLC analysis of culture supernatants. Soy (A) and SA (B) culture supernatants from wild type and Δ*cea*S2::*apr* mutant cultures were used in the analysis. Of the two Δ*cea*S2::*apr* mutants fermented, chromatograms for only one (Δ*cea*S2::*apr* #2) are shown. The wild type samples were diluted (1/5) before analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.

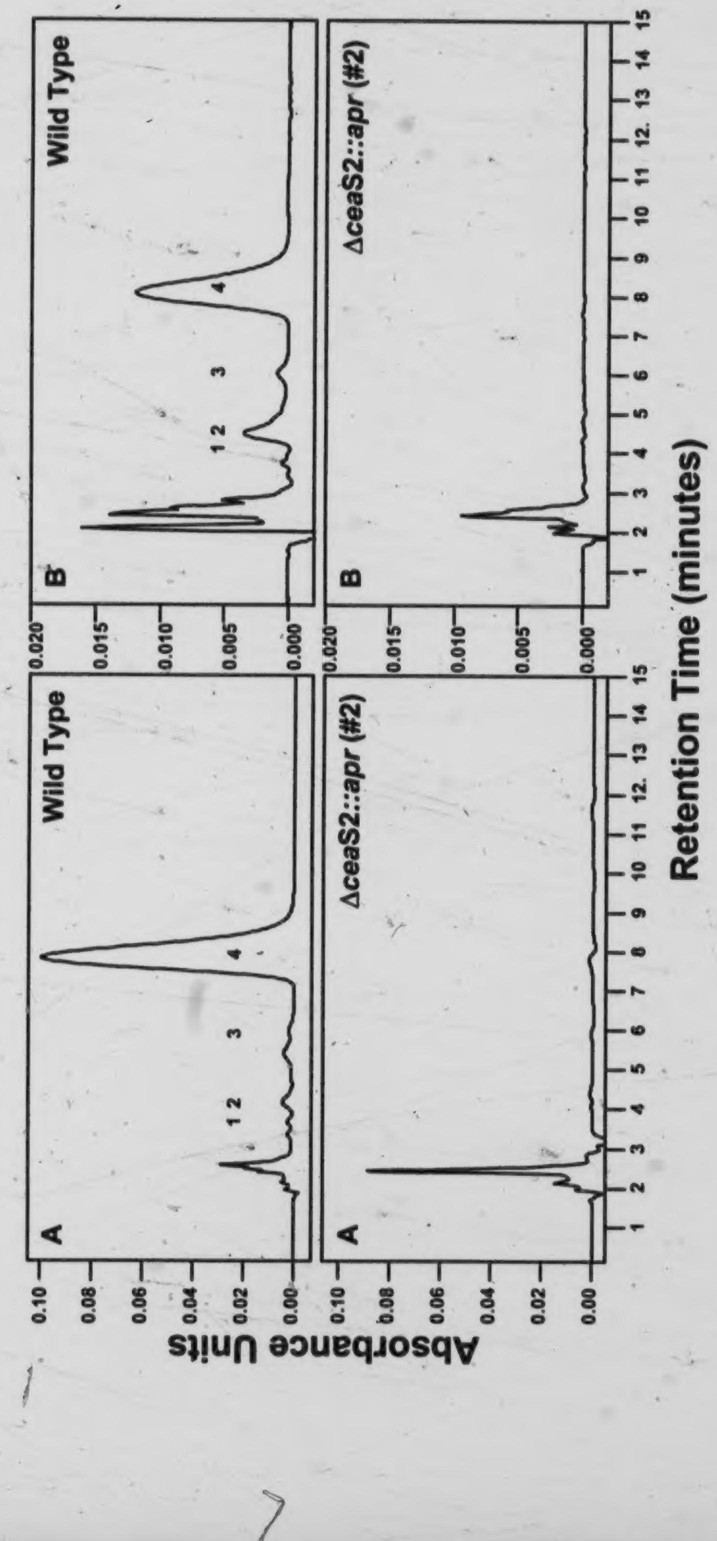

Absorbance Units
A
Wild Type
1 2
3
4
0.10
0.08
0.06
0.04
0.02
0.00
B
Wild Type
1 2
3
4
0.020
0.015
0.010
0.005
0.000
A
ΔceaS2::apr (#2)
0.10
0.08
0.06
0.04
0.02
0.00
B
ΔceaS2::apr (#2)
0.020
0.015
0.010
0.005
0.000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Retention Time (minutes)

Figure 3.1.14 Preparation of pUWL-Δ*cea*S2, the Δ*cea*S2 in-frame disruption construct. The mutagenized recombinant cosmid 12B8-AP (see Figure 3.1.10) in which *cea*S2 was deleted and replaced by the [*acc(3)IV+oriT*] disruption cassette was used in the preparation of the plasmid pUWL-Δ*cea*S2::*apr*, and then pUWL-Δ*cea*S2. The hollow circles represent the plasmid vectors and the hollow boxes represent the region of the *S. clavuligerus* chromosome surrounding *cea*S2. The filled circles represent the FLP recombination target (FRT) sites. The restriction sites and the antibiotic markers used in the preparation of the respective plasmids are also shown.

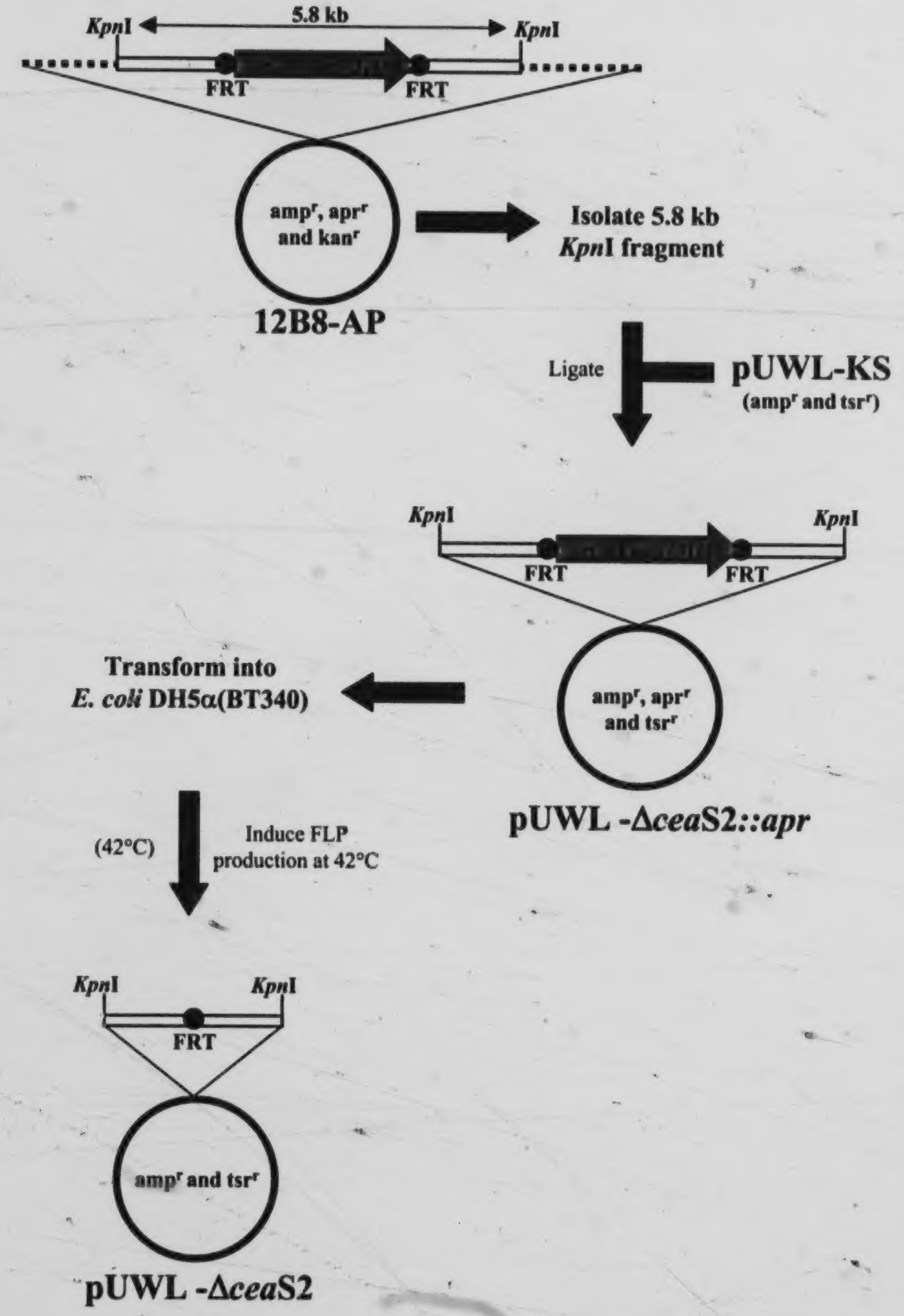
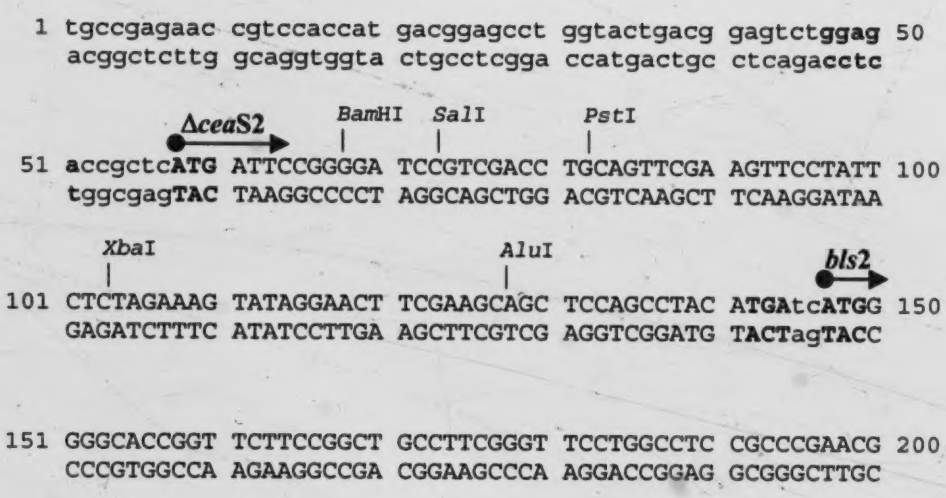

KpnI
5.8 kb
KpnI
FRT
FRT
ampʳ, aprʳ and kanʳ
12B8-AP
Isolate 5.8 kb KpnI fragment
Ligate
pUWL-KS
(ampʳ and tsrʳ)
KpnI
KpnI
FRT
FRT
ampʳ, aprʳ and tsrʳ
pUWL -ΔceaS2::apr
Transform into E. coli DH5α(BT340)
(42°C)
Induce FLP production at 42°C
KpnI
KpnI
FRT
ampʳ and tsrʳ
pUWL -ΔceaS2

Figure 3.1.15 DNA sequence of the region corresponding to the Δ*cea*S2 in-frame deletion in pUWL-Δ*cea*S2. The DNA sequence was obtained using the primer APA17 and was used to confirm that the entire *cea*S2 gene had been deleted in-frame. The 81 bp "scar" left after the deletion is shown in capital letters and so is the *bls*2 coding sequence. The Shine-Dalgarno sequence, translational start and translational stop codons for *cea*S2, and the translational start codon for *bls*2 are shown in bold letters. The restriction sites introduced into pUWL-Δ*cea*S2, which were not present in the original *cea*S2 coding sequence, are also indicated.



```
  1 tgccgagaac cgtccaccat gacggagcct ggtactgacg gagtctggag 50
    acggctcttg gcaggtggta ctgcctcgga ccatgactgc ctcagacctc

             ΔceaS2        BamHI  SalI       PstI
                             |      |          |
 51 accgctcATG ATTCCGGGGA TCCGTCGACC TGCAGTTCGA AGTTCCTATT 100
    tggcgagTAC TAAGGCCCCT AGGCAGCTGG ACGTCAAGCT TCAAGGATAA

      XbaI                       AluI                bls2
       |                           |
101 CTCTAGAAAG TATAGGAACT TCGAAGCAGC TCCAGCCTAC ATGAtcATGG 150
    GAGATCTTTC ATATCCTTGA AGCTTCGTCG AGGTCGGATG TACTagTACC

151 GGGCACCGGT TCTTCCGGCT GCCTTCGGGT TCCTGGCCTC CGCCCGAACG 200
    CCCGTGGCCA AGAAGGCCGA CGGAAGCCCA AGGACCGGAG GCGGGCTTGC
```

Δ*cea*S2::*apr*, leaving behind an in-frame deletion in place of Δ*cea*S2::*apr* to give 12B8-Δ*cea*S2 and pUWL-Δ*cea*S2, respectively (Figure 3.1.14). The preparation of the *cea*S2 in-frame deletion in pUWL-Δ*cea*S2 was confirmed by sequencing across the Δ*cea*S2 region using the primer APA17 (Figure 3.1.15). The plasmid pUWL-Δ*cea*S2 was then transformed into *E. coli* ET12567 and plasmid DNA was re-isolated, and both denatured and non-denatured plasmid DNA was then transformed into *cea*S2::*apr* (4B) mutant protoplasts. Primary transformants that were apr^r and tsr^r resistant were isolated and allowed to undergo two rounds of sporulation under nonselective conditions to isolate apr^s and tsr^s, Δ*cea*S2 mutants which arose by homologous recombination (Figure 3.1.16).

Two mutants were isolated and verified by Southern analysis to confirm that gene replacement had taken place as described in Section 3.1.2.1. Genomic DNA from the wild type and the Δ*cea*S2 mutant strains (#1 and #2) was digested with *Nco*I and was used in the analysis (Figure 3.1.17A). When blots containing bound chromosomal DNA were used, the *cea*S2-specific probe hybridized to a 2.87 kb fragment in lanes containing DNA from the wild type strain, whereas the same hybridizing fragment was not observed in lanes containing Δ*cea*S2 mutant DNA (Figure 3.1.17B). In addition, when an *apr*-specific probe was used in the analysis, none of the DNA samples hybridized to the probe (Figure 3.1.17B). This indicated that the chromosomal *apr*-disrupted copy of *cea*S2 in the 4B mutant had been replaced by the in-frame deleted copy of *cea*S2, leaving the downstream ORF intact (Figure 3.1.17A).

3.1.3.2 Phenotype of the Δ*cea*S2 mutant

Two Δ*cea*S2 mutants (#1 and #2) were fermented along with the wild type strain for 72 and 96 hours in soy fermentation medium, and culture supernatants were analyzed by HPLC for antibiotic production as described earlier. The wild type strains produced the expected levels of clavulanic acid and the 5*S* clavams in soy medium; whereas none of the Δ*cea*S2 mutants produced the respective metabolites under similar conditions (Figure 3.1.18).

3.1.3.3 Complementation of the Δ*cea*S2 mutant

Figure 3.1.16 Schematic of the process leading to the isolation of the Δ*cea*S2 in-frame deletion mutant using the plasmid pUWL-Δ*cea*S2. The plasmid pUWL-Δ*cea*S2 in which *cea*S2 has been deleted in-frame (see Figure 3.1.14) was introduced into *S. clavuligerus cea*S2*::apr* (4B) mutant protoplasts by transformation. The gray and black arrows represent *cea*S2 and the apramycin resistance gene cassette (*apr*) respectively, with the arrowheads indicating the direction of transcription. The hollow boxes represent the region of the *S. clavuligerus* chromosome surrounding *cea*S2 and the filled circles represent the FLP recombination target (FRT) sites, which are part of the in-frame deletion. Recombination between the homologous regions of pUWL-Δ*cea*S2 and the *S. clavuligerus cea*S2*::apr* (4B) mutant chromosome led to the replacement of the chromosomal copy of *cea*S2*::apr*, by the Δ*cea*S2 in-frame deletion. This was followed by the loss of the plasmid resulting in the apr^s and tsr^s Δ*cea*S2 in-frame deletion mutants.

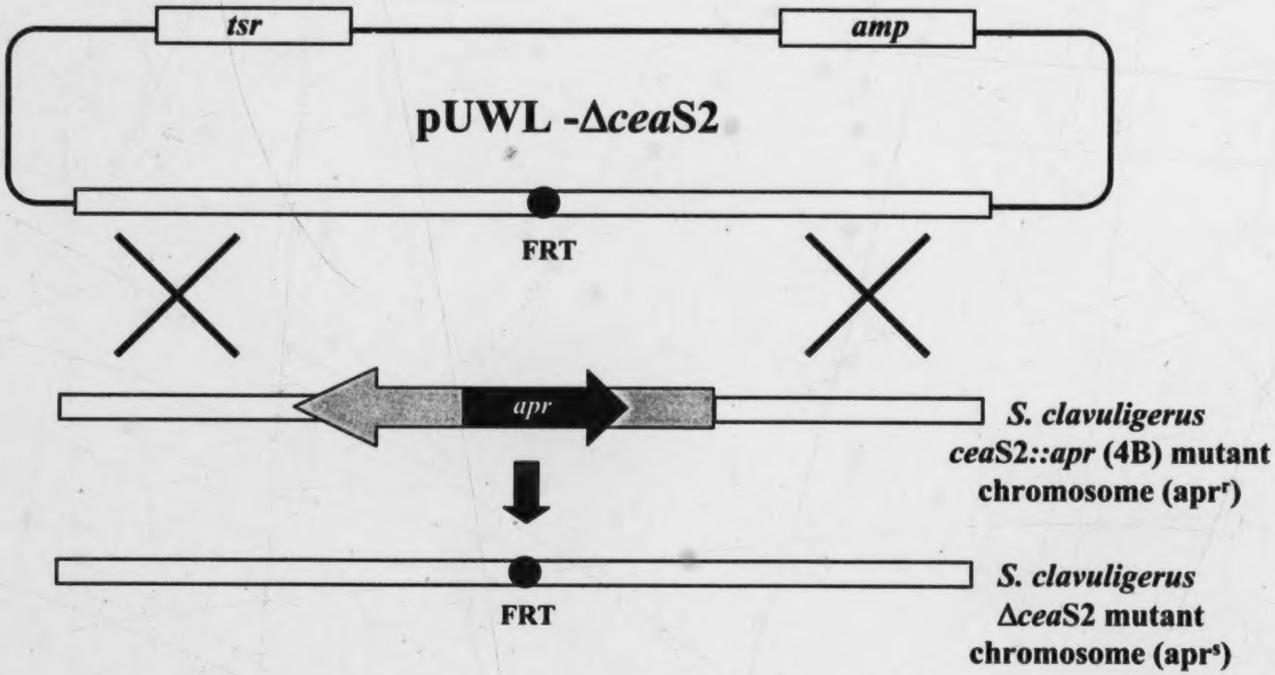
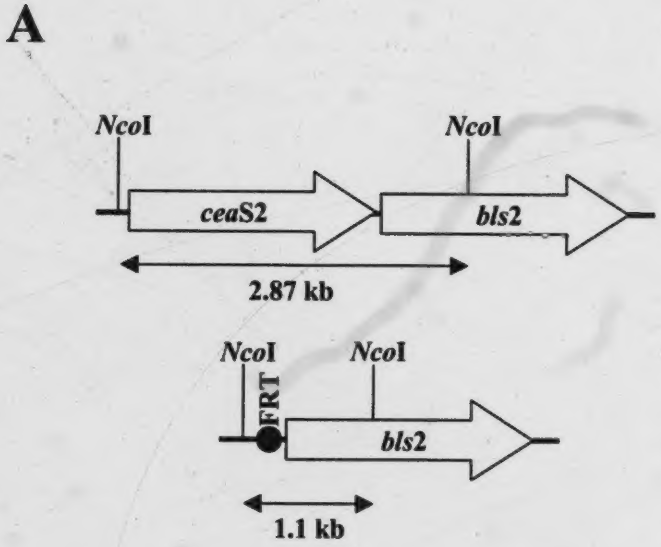

tsr
amp
pUWL -ΔceaS2
FRT
apr
S. clavuligerus ceaS2::apr (4B) mutant chromosome (apr^r)
FRT
S. clavuligerus ΔceaS2 mutant chromosome (apr^s)

Figure 3.1.17 Southern analysis of the *S. clavuligerus* Δ*cea*S2 mutant. (A) Diagrammatic representation of the *cea*S2 region of the *S. clavuligerus* chromosome in the wild type and the Δ*cea*S2 mutants. The open arrows represent *cea*S2 and *bls*2 with the direction of transcription represented by the arrowheads. The filled circle represents the in-frame deletion in the Δ*cea*S2 mutants and the fine lines represent the rest of the *S. clavuligerus* chromosome. (B) Southern analysis of *Nco*I digested genomic DNA from *S. clavuligerus* wild type (WT) and Δ*cea*S2 (#1 and #2) mutant strains. After processing, DNA from the respective strains was probed using a *cea*S2-specific and an *acc(3)IV+oriT*-specific probe, as the parental strain of the Δ*cea*S2 mutant was the *cea*S2*::apr* (4B). The additional restriction fragments that weakly hybridize to the *cea*S2-specific probe arise due to the presence of extra DNA in the probe preparation, corresponding to the region surrounding *cea*S2. In addition, strain #1 shows the presence of a large restriction fragment that strongly hybridizes to the *cea*S2-specific probe, which cannot be accounted for.


A
NcoI
NcoI
ceaS2
bls2
2.87 kb
NcoI
NcoI
FRT
bls2
1.1 kb

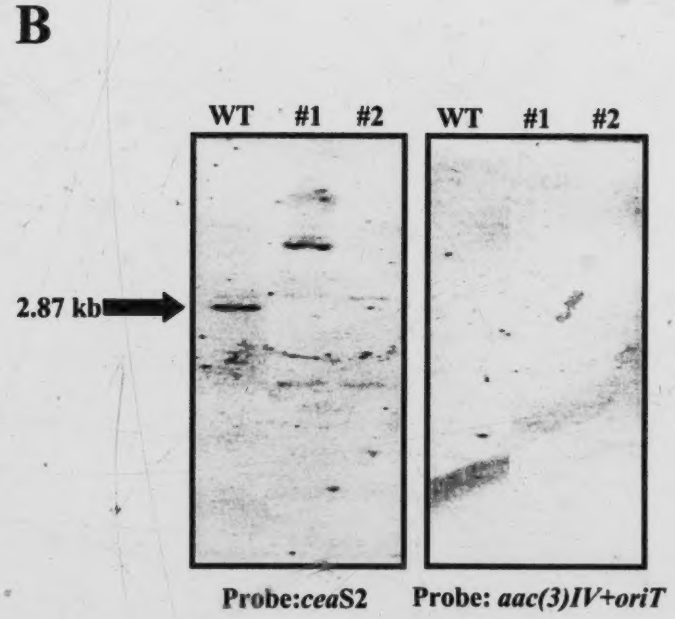
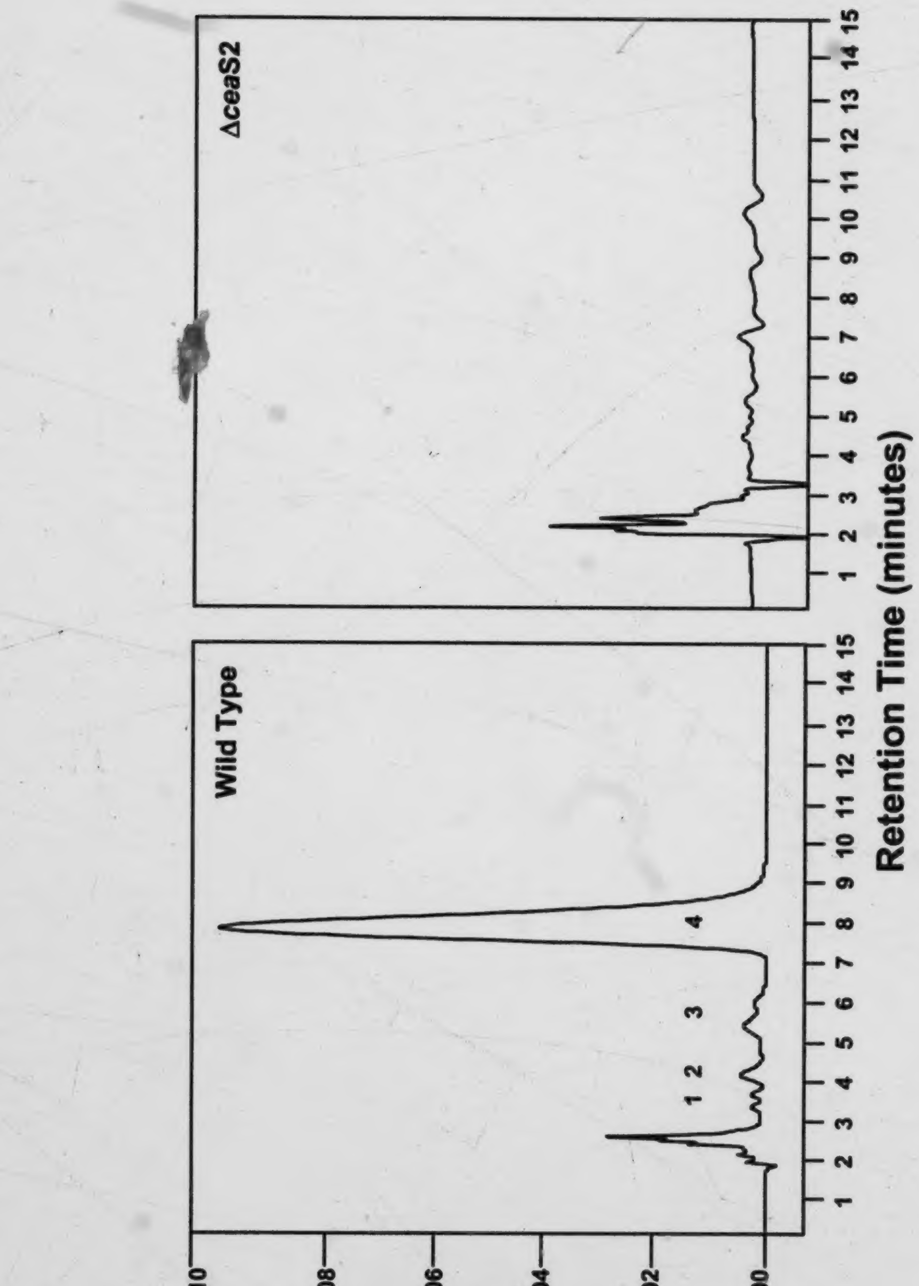

B
WT
#1
#2
WT
#1
#2
2.87 kb
Probe:ceaS2
Probe: aac(3)IV+oriT

Figure 3.1.18 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* Δ*cea*S2 in-frame deletion mutant. Soy culture supernatants (96 hour) from the wild type and Δ*cea*S2 in-frame deletion mutants were analyzed by HPLC. Of the two Δ*cea*S2 mutants fermented, the chromatogram for only one (Δ*cea*S2 #2) is shown. The wild type sample was diluted (1/5) before analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.


Absorbance Units
0.10
0.08
0.06
0.04
0.02
0.00
Wild Type
1
2
3
4
ΔceaS2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Retention Time (minutes)

To determine if the Δ*cea*S2 mutants were blocked in clavulanic acid and 5*S* clavam production due to polar effects, the Δ*cea*S2 #2 mutant was complemented using the *cea*S2 complementation construct pSET152-5.3 described earlier (Section 3.1.1.3). As well, a mutant strain harboring the plasmid vector pSET152 without any insert was also isolated and analyzed as a control.

The wild type, Δ*cea*S2 mutant, Δ*cea*S2 complemented with pSET152-5.3 and the Δ*cea*S2 control strain containing the plasmid pSET152 were fermented in soy medium and analyzed as described earlier. Again, under the conditions tested, only the wild type strain produced detectable levels of clavulanic acid and the 5*S* clavams. This indicated that the plasmid pSET152-5.3 could not complement the Δ*cea*S2 mutant.

3.2 Isolation and DNA sequence of the *cea*S paralogue

Since the *cea*S2-Fs mutants produced some clavulanic acid and 5*S* clavams, even though the amounts were very small, this indicated that some other protein could functionally replace the *cea*S2 gene product. A similar situation was observed earlier in the case of *cas*2 and *pah*2, although mutants of either *cas*2 or *pah*2 still produced considerable amounts of clavulanic acid and the 5*S* clavams (Table 1.2), and paralogues of these genes have been isolated and characterized. This led us to believe that a second copy of *cea*S existed in *S. clavuligerus*; therefore, attempts were made to isolate the *cea*S paralogue. Note: portions of this section have been published as part of the manuscript by Tahlan *et al.* (2004c).

3.2.1 Isolation and sequencing of the *cea*S paralogue

While performing Southern analysis to confirm gene replacement in the *cea*S2-Fs mutant strains, multiple bands were observed to cross-hybridize to the *cea*S2-specific probe in all of the lanes (Figure 3.1.3). The cross-hybridizing bands were postulated to represent fragments of a gene, which shared sequence similarity with *cea*S2. To investigate the possibility that this could be the second *cea*S gene in *S. clavuligerus*, chromosomal DNA from wild type *S. clavuligerus* was digested with *Eco*RI and was analyzed by Southern hybridization using a *cea*S-specific probe. The 855 bp *Eco*RI-*Not*I fragment from pCAD2-3, which contained the 5′ region of *cea*S2 and some upstream coding sequence, was again used as the *cea*S2-specific probe. A DNA fragment of 12 kb hybridized strongly with the probe while a 2.85 kb fragment gave a weaker hybridization signal (Figure 3.2.1A). From previous studies, it was known that the *cea*S2 gene in the clavulanic acid gene cluster is carried on a 12 kb *Eco*RI fragment (Jensen *et al*., 2000) (Figure 1.4.2). The 2.85 kb *Eco*RI fragment was therefore postulated to encode a second copy of *cea*S, or the putative *cea*S paralogue.

Previous DNA sequence analyses had not located any putative paralogues in the regions flanking *cas*1 (Mosher *et al*., 1999) and *pah*1; the paralogue of *pah*2 was ultimately located in a separate region of the *S. clavuligerus* chromosome (Jensen *et al*., 2004b). Therefore, it seemed possible that other paralogues, if they existed, might be clustered with *pah*1. Cosmids 14E10 and 6G9, previously shown to carry *pah*1 and flanking regions of the chromosome (Jensen *et al*., 2004b), were digested with *Eco*RI and were subjected to Southern analysis using the *cea*S2-specific probe. A 2.85 kb hybridizing fragment was observed in both of the cosmids (Figure 3.2.1B), and the fragment was subcloned for further study giving the plasmid p2.8-18. The 2.85 kb *Eco*RI insert from plasmid p2.8-18 was partially sequenced using custom primers (Table 2.4), and by sequencing p2.8-18 and other subclones (Table 2.3), using universal and reverse primers. Sequence analysis indicated the presence of an incomplete ORF showing similarity to the 5' end of *cea*S2, located at one end of this fragment. At the same time, during a search for the *bls* paralogue in *S. clavuligerus*, a 5.7 kb *Eco*RI fragment present in cosmid 6G9 but not in 14E10, which hybridized to a *bls*-specific probe, was subcloned and partially sequenced (Tahlan *et al*., 2004c). At one end of this fragment the 3′ end of the *cea*S2-like ORF was found. All DNA sequence information was confirmed on both strands, and sequence information was obtained directly from the cosmids 14E10 and 6G9 using custom primers, to cross all junctions of subclones, in order to ensure that no small fragments were lost during subcloning. Genes in streptomycetes and other organisms with high G+C content in their DNA, exhibit a bias in codon usage, resulting in very high G+C levels at the third letter position of each codon. Therefore, the FramePlot program predicted the presence of a complete 1668 bp ORF with a 3rd-Letter G+C content of 95.5% (Figure 3.2.2), which is consistent with the codon usage in *Streptomyces* ORFs (Bibb *et al*., 1984). The nucleotide sequence of this

Figure 3.2.1 Isolation of the *cea*S paralogue (*cea*S1). (A) Southern analysis of genomic DNA from wild type *S. clavuligerus* after *Eco*RI digestion. The fractionated and blotted genomic DNA was probed using a *cea*S2-specific probe. The hybridizing 12 kb fragment known to contain *cea*S2, and the cross-hybridizing 2.85 kb fragment, thought to contain the putative *cea*S paralogue are shown. (B) Southern analysis of cosmids 6G9 and 14E10 using a *cea*S2-specific probe. Lane1; *cea*S2 probe (control), Lanes 2 and 3, *Eco*RI digested 6G9 and 14E10 respectively.

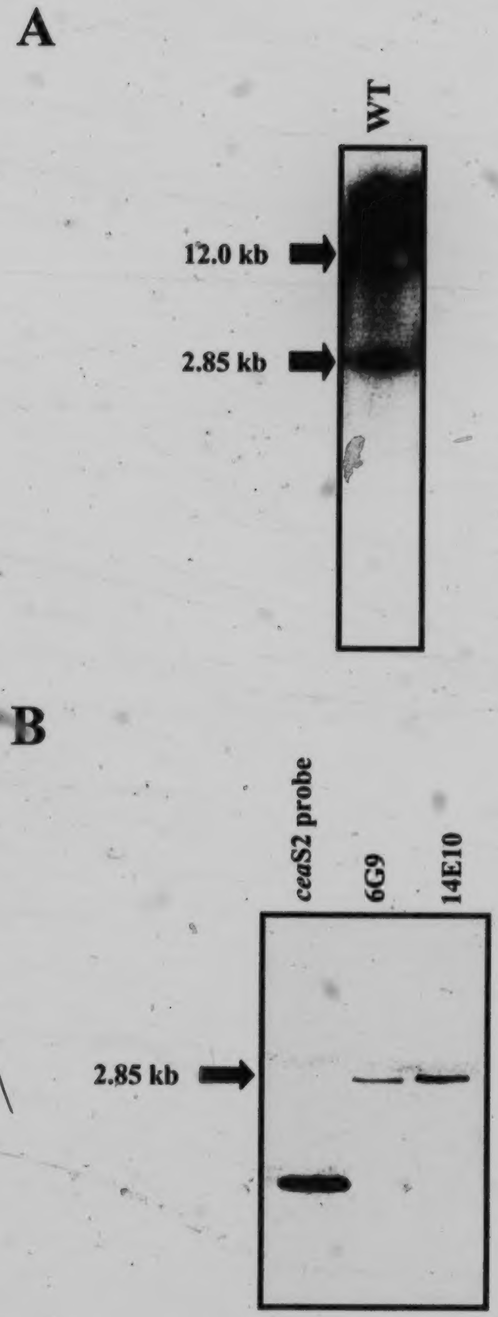
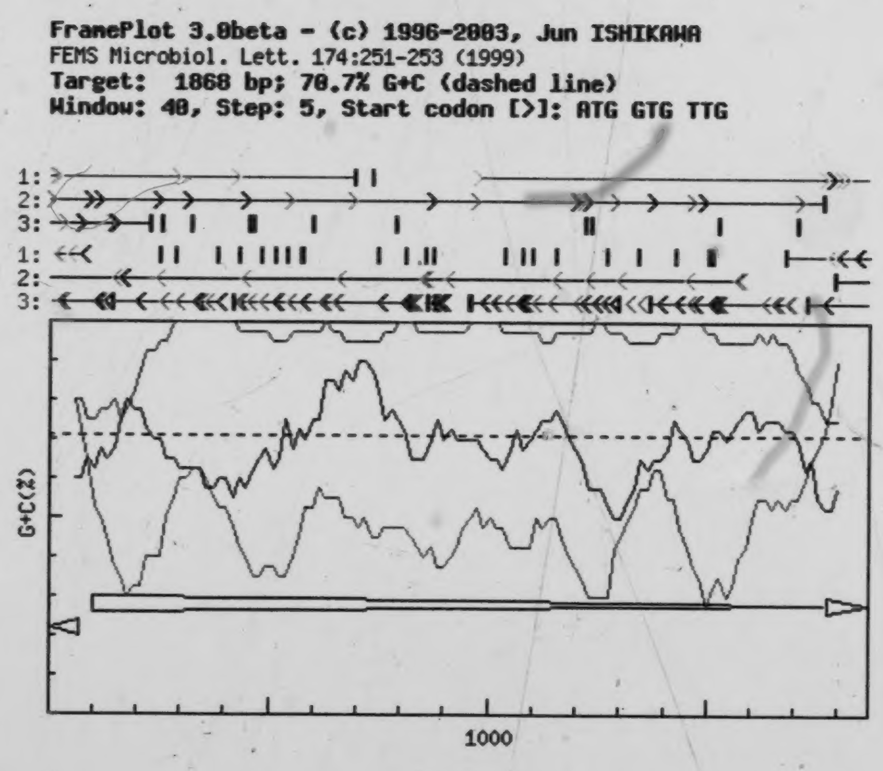

A
WT
12.0 kb
2.85 kb
B
ceaS2 probe
6G9
14E10
2.85 kb

Figure 3.2.2 Analysis of the *cea*S1 coding sequence using FramePlot. The complete *cea*S1 open reading frame (ORF) was isolated and identified by partially sequencing the ends of the plasmids p2.8-18 and p5.7, in addition to obtaining DNA sequence directly from the cosmids 14E10 and 6G9. The sequence data obtained was analyzed using FramePlot, which identifies ORFs by measuring G+C percentages as a function of codon position. Genes in organisms with high G+C content in their DNA exhibit a bias in codon usage, resulting in very high G+C levels at the third letter position of each codon. The lines above the graphs represent the six possible reading frames, the arrowheads represent potential start codons and the vertical bars represent potential stop codons.



FramePlot 3.0beta - (c) 1996-2003, Jun ISHIKAWA
FEMS Microbiol. Lett. 174:251-253 (1999)
Target: 1868 bp; 70.7% G+C (dashed line)
Window: 40, Step: 5, Start codon [>]: ATG GTG TTG

newly isolated ORF shared 73% identity with that of *cea*S2. The newly isolated *cea*S-like ORF was called *cea*S1 (Figure 3.2.3), to distinguish it from *cea*S2, which is located in the clavulanic acid gene cluster. Upstream of *cea*S1 (278 bp), an incomplete ORF showing 34% identity (44% similarity) at the amino acid level to different glycine and serine hydroxymethyltransferases, was partially sequenced and was found to be oriented in the direction opposite to *cea*S1 transcription (Figures 3.2.3 and 3.2.4). Results from other related studies (Jensen *et al.*, 2004b; Tahlan *et al.*, 2004c), showed that second, or paralogous copies of *bls*, *pah* and *oat* (now referred to as *bls*1, *pah*1 and *oat*1) were all located downstream of *cea*S1 (Figure 3.2.4). The gene cluster comprising *cea*S1, *bls*1, *pah*1 and *oat*1 was therefore referred to as the paralogue gene cluster (Figure 3.2.4) to distinguish it from the clavulanic acid gene cluster (Figure 1.4.2), which encodes *cea*S2, *bls*2, *pah*2, *cas*2, and *oat*2 in addition to other genes involved in clavulanic acid biosynthesis.

3.2.2 Protein sequence analysis

The predicted amino acid sequence of CeaS1 showed 66% identity (77% similarity) to CeaS2 from the clavulanic acid gene cluster. When *cea*S2 was first sequenced, it showed striking similarity to genes encoding acetohydroxyacid synthases (AHAS) based on observed homologies, and conservation of five of the eight amino acids forming the active center of AHAS (Perez-Redondo *et al.*, 1999). Subsequently, CeaS2 was shown to catalyze the thiamine pyrophosphate (TPP)-dependent condensation of glyceraldehyde-3-phosphate and L-arginine to form carboxyethylarginine (Figure 1.4.1), the first reaction in the clavulanic acid-clavam biosynthetic pathways (Khaleeli *et al.*, 1999). The predicted CeaS1 protein also showed varying levels of similarity to AHAS large subunits and other TPP requiring proteins from various organisms (Figure 3.2.5). The five amino acids associated with the active centers of AHAS enzymes and which are found in CeaS2 (^{57}E,^{121}Q,^{464}G,^{489}N and ^{594}G) are also conserved in CeaS1 (^{48}E,^{112}Q,^{446}G,^{471}N and ^{476}G). Since CeaS2 utilizes TPP, it also contains the TPP binding motif ($^{446}\mathrm{I\underline{G}AQMAR\underline{P}DQPTFLIA\underline{GDG}G}^{465}$), and a similar TPP binding motif ($^{428}\mathrm{M\underline{A}AQIAR\underline{P}GEPVFLIA\underline{GDG}G}^{447}$) is also present in the predicted CeaS1 protein.

Figure 3.2.3 DNA sequence of *cea*S1. The complete DNA sequence of *cea*S1 and downstream ORFs (up to *oat*1) has been deposited in GenBank under the accession number AY426768 (Tahlan *et al.* 2004c). Sequences of ORFs are shown in upper case whereas intergenic regions are shown in lower case. Important restriction sites and the plasmids containing the respective portions of *cea*S1 are also indicated. The predicted translational start and stop codons, and Shine-Dalgarno sequences are shown in bold letters and the arrowheads indicate the direction of transcription of the respective genes. The partial DNA sequence of *orf*A (This study) and *bls*1 (Tahlan *et al.* 2004c) are shown for reference only, and do not represent the complete DNA inserts in the plasmids p2.8-18 and p5.7, respectively.

```
  1 CCGGCGTGGC GCGCCCCGAA GAAGCGCTTG GCCAGGTCCT CGCAGTACGC 50
    GGCCGCACCG CGCGGGGCTT CTTCGCGAAC CGGTCCAGGA GCGTCATGCG

 51 GTACACCTCG GCCAGCCCCG TGACATCGCC GTAGACGGCC ACCGGGGTCT 100
    CATGTGGAGC CGGTCGGGGC ACTGTAGCGG CATCTGCCGG TGGCCCCAGA

101 CGGAGTACGG ATAGCGGTTG ACGGCGTCGG TGGCCAGCGC GGCACTGGCG 150
    GCCTCATGCC TATCGCCAAC TGCCGCAGCC ACCGGTCGCG CCGTGACCGC

151 CGCGGCGACA GCCGGTTCTC GATGGGGAAA AGATTCAGGC TGGGCTTGCG 200
    GCGCCGCTGT CGGCCAAGAG CTACCCCTTT TCTAAGTCCG ACCCGAACGC
                           NcoI
                            |
201 CTCCAACGCG GCCAGGACAT CCATggattc tcctcggatc ggtggacgga 250
    GAGGTTGCGC CGGTCCTGTA GGTAcctaag aggagcctag ccacctgcct
                            orfA

251 aggggacgga atcggatgga gaccgtcggc gccggacgga tttcacgaaa 300
    tcccctgcct tagcctacct ctggcagccg cggcctgcct aaagtgcttt

301 agcgccacga ggacgatgga cgccatgcct tgaacgcgcc ttgcccattc 350
    tcgcggtgct cctgctacct gcggtacgga acttgcgcgg aacgggtaag

351 tgtccaagcc gggcgtccga ccgccgccaa gaggcgatcc gcacagtggc 400
    acaggttcgg cccgcaggct ggcggcggtt ctccgctagg cgtgtcaccg

401 aacgccacca cccacggtgt gcgccacccg gtgcgcaccg cccggtgccg 450
    ttgcggtggt gggtgccaca cgcggtgggc cacgcgtggc gggccacggc
                                                    NcoI
                                                     |
451 cgttttccgg ggcgccatcc cggcgcccgt ccgatgcgaa ggagatctcc 500
    gcaaaaggcc ccgcggtagg gccgcgggca ggctacgctt cctctagagg

      ceaS1
501 ATGGCCACCA CGACCGCGAA AGCCATGCTG GAACGTCTTC ACCAGTACGG 550
    TACCGGTGGT GCTGGCGCTT TCGGTACGAC CTTGCAGAAG TGGTCATGCC

551 TGTCGACCAT GTATTCGGCG TCGTCGGCCG GGAGGCGTCC GCCATTCTCT 600
    ACAGCTGGTA CATAAGCCGC AGCAGCCGGC CCTCCGCAGG CGGTAAGAGA

601 TCGACGAGGT CGAAGGACTC GACTTCGTCC TGACCCGGCA CGAGTTCACC 650
    AGCTGCTCCA GCTTCCTGAG CTGAAGCAGG ACTGGGCCGT GCTCAAGTGG

651 GCCGGGGTGA TGGCGGACGT CCTCGCCCGG ATCACCAACC GCCCCCAGGC 700
    CGGCCCCACT ACCGCCTGCA GGAGCGGGCC TAGTGGTTGG CGGGGGTCCG
```

```
                           ApaI
                           |
 701 GTGCTTCGCG ACCCTGGGCC CCGGCATGAC CAACCTGGCC ACCGGCGTCG 750
     CACGAAGCGC TGGGACCCGG GGCCGTACTG GTTGGACCGG TGGCCGCAGC

 751 CCACCTCCGC CCTGGACCGC AGCTCGGTCA TCGCGCTGGC CGCGCAGTCC 800
     GGTGGAGGCG GGACCTGGCG TCGAGCCAGT AGCGCGACCG GCGCGTCAGG

 801 GAGTCGTACG ACTGCTACCC CAACGTCACC CACCAGTGCC TGGACAGCAC 850
     CTCAGCATGC TGACGATGGG GTTGCAGTGG GTGGTCACGG ACCTGTCGTG
                    ApaI
                    |
 851 CGCCGTGATG GGCCCGCTGA CCAAGTTCAG CGTCCAGCTC GAACGCGGCG 900
     GCGGCACTAC CCGGGCGACT GGTTCAAGTC GCAGGTCGAG CTTGCGCCGC

 901 AGGACATCGT CAACCTCGTC GACAGCGCCG TCCTCAACAG CCGGATCGAG 950
     TCCTGTAGCA GTTGGAGCAG CTGTCGCGGC AGGAGTTGTC GGCCTAGCTC

 951 CCCGTGGGTC CCAGCTTCAT CAGCCTGCCG GTCGACCTCC TCGGCGCCGA 1000
     GGGCACCCAG GGTCGAAGTA GTCGGACGGC CAGCTGGAGG AGCCGCGGCT

1001 GCTGAACGGC ACCCCCACCG ACGCCCCCCT GGTCCGGGCC ACCGCCACCC 1050
     CGACTTGCCG TGGGGGTGGC TGCGGGGGGA CCAGGCCCGG TGGCGGTGGG

1051 ACGCCCTGGA CGCCGACTGG CGCGCCCGCC TCGACGAGGC CGCTGAGCTG 1100
     TGCGGGACCT GCGGCTGACC GCGCGGGCGG AGCTGCTCCG GCGACTCGAC

1101 GTGCGCGAGG CCGAGAACCC CCTCCTCGTC GTCGGTAGCG CCGTCATCCG 1150
     CACGCGCTCC GGCTCTTGGG GGAGGAGCAG CAGCCATCGC GGCAGTAGGC

1151 CGCCGGGGCC GTCGACGCCC TGCGCGCCCT CGCCGAGCGG CTGAACATCC 1200
     GCGGCCCCGG CAGCTGCGGG ACGCGCGGGA GCGGCTCGCC GACTTGTAGG

1201 CCGTCGTCAC CACCTACACC GCCAAGGGCG TCCTGCCGCA CGACCACCCG 1250
     GGCAGCAGTG GTGGATGTGG CGGTTCCCGC AGGACGGCGT GCTGGTGGGC

1251 CTCAACTACG GCGCCATCAG CGGCTACATG GACGGCATTC TCGGCCACCC 1300
     GAGTTGATGC CGCGGTAGTC GCCGATGTAC CTGCCGTAAG AGCCGGTGGG

1301 GGCCCTCGAC GAGATCTTCG GCCCCGCCGA CCTCCTCCTG GCGATCGGCT 1350
     CCGGGAGCTG CTCTAGAAGC CGGGGCGGCT GGAGGAGGAC CGCTAGCCGA

1351 ACGACTACGC CGAGGACCTG CGCCCCTCCA TGTGGACGCG GGGCCGGGCC 1400
     TGCTGATGCG GCTCCTGGAC GCGGGGAGGT ACACCTGCGC CCCGGCCCGG
```

```
1401 AAGACCACGG TCCGGGTCGC CCCCGAGGTC AACCCGATCC CGGAGCTGTT 1450
     TTCTGGTGCC AGGCCCAGCG GGGGCTCCAG TTGGGCTAGG GCCTCGACAA
                                               EcoRI
                           ◄──────── p2.8-18|
1451 CCGCGCCGAC ATCGACATCG TCACCAACGT CGCCGAATTC GTCACCGCGC 1500
     GGCGCGGCTG TAGCTGTAGC AGTGGTTGCA GCGGCTTAAG CAGTGGCGCG

1501 TCGACGACGC GACCTCGGGC CTCGCCCCCA AGACCCGGCA CGACCTCAGC 1550
     AGCTGCTGCG CTGGAGCCCG GAGCGGGGGT TCTGGGCCGT GCTGGAGTCG
                          EcoRI
                          |p5.7──────────────►
1551 GCCCTGCGCG CCCGCGTCGC CGAATTCCTC GCCGACCCCA CCGAGTACGA 1600
     CGGGACGCGC GGGCGCAGCG GCTTAAGGAG CGGCTGGGGT GGCTCATGCT

1601 GGACGGCATG CGGGTCCACC AGGTGATCGA CTGCATGAAC TCCGTCCTCG 1650
     CCTGCCGTAC GCCCAGGTGG TCCACTAGCT GACGTACTTG AGGCAGGAGC

1651 ACAACGGCAC CTTCGTCAGC GACATCGGCT TCTTCCGCCA CTACGGCGTG 1700
     TGTTGCCGTG GAAGCAGTCG CTGTAGCCGA AGAAGGCGGT GATGCCGCAC

1701 CTCTTCGCCA AGTCCGACCA GCCGTACGGA TTCCTCACCT CCGCGGGCTG 1750
     GAGAAGCGGT TCAGGCTGGT CGGCATGCCT AAGGAGTGGA GGCGCCCGAC
                                      NcoI
                                      |
1751 CTCCAGCTTC GGCTACGGAC TGCCCGCCGC CATGGCCGCC CAGATCGCCC 1800
     GAGGTCGAAG CCGATGCCTG ACGGGCGGCG GTACCGGCGG GTCTAGCGGG

1801 GGCCCGGCGA GCCCGTCTTC CTCATCGCGG GCGACGGCGG CTTCCACTCC 1850
     CCGGGCCGCT CGGGCAGAAG GAGTAGCGCC CGCTGCCGCC GAAGGTGAGG

1851 AACAGCGCCG ACATCGAGAC GGCCGTGCGC CTGGGCCTGC CGATCGTCAT 1900
     TTGTCGCGGC TGTAGCTCTG CCGGCACGCG GACCCGGACG GCTAGCAGTA

1901 GGTCGTCGTC AACAACGACC GCAACGGCCT GATCGAGCTG TACCAGAACC 1950
     CCAGCAGCAG TTGTTGCTGG CGTTGCCGGA CTAGCTCGAC ATGGTCTTGG

1951 TCGGACACCA GCGCTCCCAC GCCCCCGCCG TCGGCTTCGG AAGCGTCGAC 2000
     AGCCTGTGGT CGCGAGGGTG CGGGGGCGGC AGCCGAAGCC TTCGCAGCTG

2001 TTCGTCCAGC TCGCCGAGGC CAACGGCTGC GAGGCCGTCC GCGCCACCGA 2050
     AAGCAGGTCG AGCGGCTCCG GTTGCCGACG CTCCGGCAGG CGCGGTGGCT

2051 CCGCACCTCG CTGCTCGCCG CCCTCACCAA GGGCGCCGGA CTCGGCCGCC 2100
     GGCGTGGAGC GACGAGCGGC GGGAGTGGTT CCCGCGGCCT GAGCCGGCGG
```

*Kpn*I
2101 CGTTCCTGAT CGAGGTACCG GTGGCCTACG ACTTCCAGTC CGGCGGTTTC 2150
GCAAGGACTA GCTCCATGGC CACCGGATGC TGAAGGTCAG GCCGCCAAAG

● ***bls*1** →
2151 GCCGCCCTGG CCATC**TGA**cc accgt**ggagg** acgagttcac c**ATG**CCCGGC 2200
CGGCGGGACC GGTAG**ACT**gg tggc**acctcc** tgctcaagtg g**TAC**GGGCCG

2201 CCCGACCTCG TGTACGGATT CCGGGTGCGC ATCGGCACCG AGGGCCGCCC 2250
GGGCTGGAGC ACATGCCTAA GGCCCACGCG TAGCCGTGGC TCCCGGCGGG

2251 CGGCGGCGGC CCCGGCGGTC ACTCCGAACC CGGCAGCGCA CCCCGCTTCG 2300
GCCGCCGCCG GGGCCGCCAG TGAGGCTTGG GCCGTCGCGT GGGGCGAAGC

2301 CCGTCCGCGG GACCCATGTC CCCGTGCACG ACGGCACCGC GTACCCGCTC 2350
GGCAGGCGCC CTGGGTACAG GGGCACGTGC TGCCGTGGCG CATGGGCGAG

2351 TGGAGCGGAA CGGCCGTGAC CCTGGGCCGT CCGCCCGTCC TGGTCGCCGA 2400
ACCTCGCCTT GCCGGCACTG GGACCCGGCA GGCGGGCAGG ACCAGCGGCT

2401 CGGCCAGGTC CGGCTGCTCC TGGCGGGCGA GCTGTACAAC CGCGCCGAGC 2450
GCCGGTCCAG GCCGACGAGG ACCGCCCGCT CGACATGTTG GCGCGGCTCG

2451 TGACCGGAGC GCTCGGCGGC TCCTCTGCCG CCCTCGGCGA CGCCGAACTG 2500
ACTGGCCTCG CGAGCCGCCG AGGAGACGGC GGGAGCCGCT GCGGCTTGAC

*Apa*I
2501 CTGCTGGCCG CCTGGCGGCG CTGGGGCCCC GGGGCCTTCC GGCTCCTGAA 2550
GACGACCGGC GGACCGCCGC GACCCCGGGG CCCCGGAAGG CCGAGGACTT

2551 CGGACGGTTC GCCGCACTGC TCACCGACGC CTCCACCGGC GCGACCGTCG 2600
GCCTGCCAAG CGGCGTGACG AGTGGCTGCG GAGGTGGCCG CGCTGGCAGC

*Kpn*I
2601 CGGCCACCGA CCACGCCGGT TCGGTACCGC TGTGGCTGCG CGCCGACGTG 2650
GCCGGTGGCT GGTGCGGCCA AGCCATGGCG ACACCGACGC GCGGCTGCAC

2651 ACGGGGCTGA GCGCCGCC 2668
TGCCCCGACT CGCGGCGG

Figure 3.2.4 Diagrammatic representation of the genes from the paralogue gene cluster present on the insert in the recombinant cosmid 14E10. Only restriction sites referred to in the text are shown. *Eco*RI* denotes an *Eco*RI site arising from the multiple cloning site of the cosmid vector, pWE15 (in the recombinant cosmid 6G9, which was also used for sequence determination and contains an insert that begins with *c6p*, which is different from 14E10). The gray bar represents the sequenced portion of the DNA insert in the recombinant cosmid 14E10 and a more detailed diagram of this region is also included. The arrows represent the respective genes with the arrowheads indicating the direction of transcription. (diagram is not to scale).

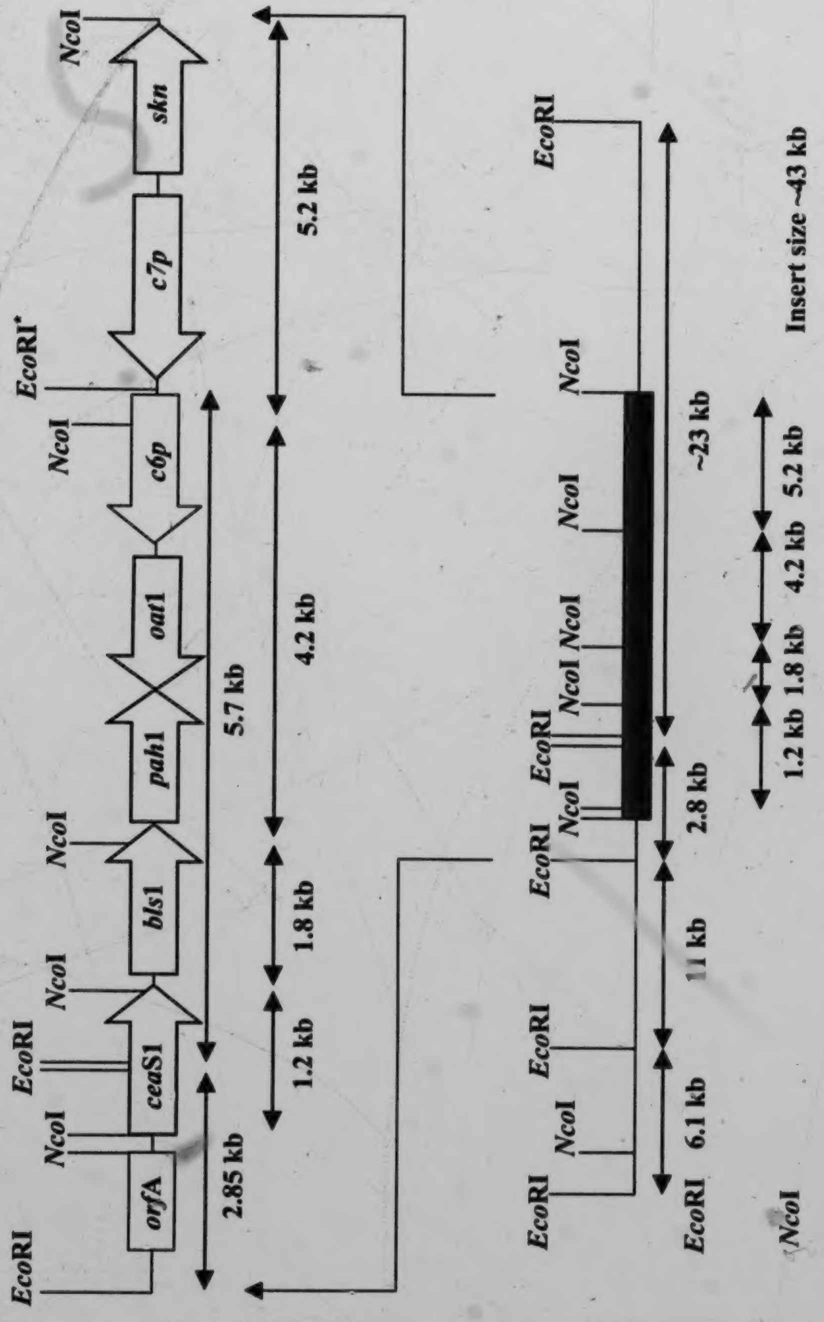
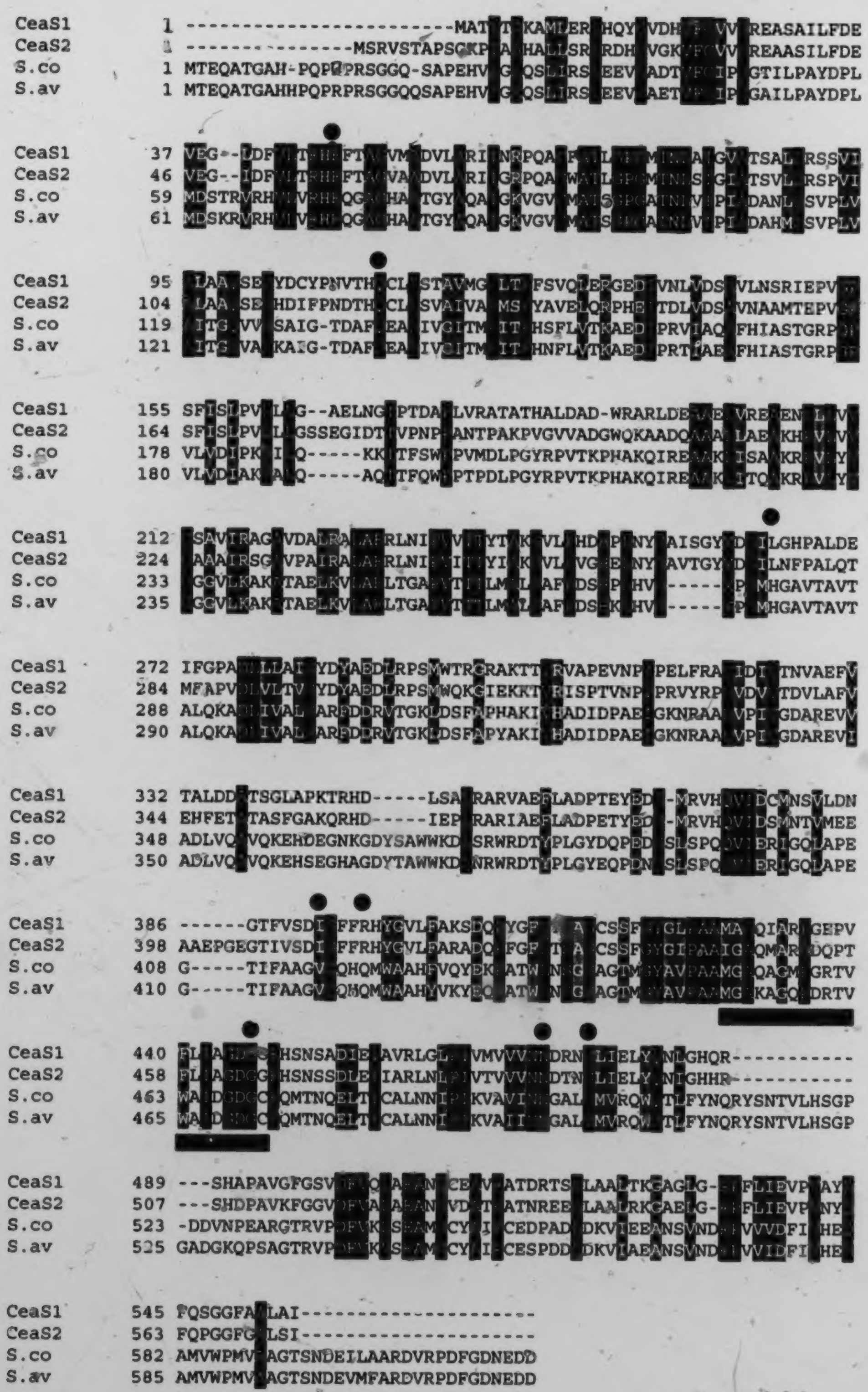

EcoRI
orfA
NcoI
ceaS1
EcoRI
NcoI
bls1
NcoI
pah1
oat1
c6p
NcoI
EcoRI*
c7p
skn
NcoI
2.85 kb
5.7 kb
1.2 kb
1.8 kb
4.2 kb
5.2 kb
EcoRI
NcoI
EcoRI
EcoRI
NcoI
EcoRI
NcoI
NcoI
NcoI
NcoI
NcoI
EcoRI
EcoRI
6.1 kb
11 kb
2.8 kb
~23 kb
1.2 kb
1.8 kb
4.2 kb
5.2 kb
Insert size ~43 kb
NcoI

Figure 3.2.5 Alignment of CeaS and acetolactate synthase/acetohydroxy synthase (AHAS) proteins from other *Streptomyces*. Conserved amino acid residues that are shown on black background, similar residues are shown on gray background and dissimilar residues are printed on white background. The filled circles represent eight residues conserved in the active sites of all AHAS enzymes and the filled box indicates the thiamine pyrophosphate (TPP) binding motif. The proteins along with their source and GenBank accession numbers are as follows: CeaS1, putative carboxyethyl arginine synthase, *Streptomyces clavuligerus*, (gi|37951558); CeaS2, carboxyethyl arginine synthase, *Streptomyces clavuligerus*, (gi|39654990); S.co, acetolactate synthase, *Streptomyces coelicolor* A3(2), (gi|21223868) and S.av, putative acetolactate synthase large subunit, *Streptomyces avermitilis* MA-4680, (gi|29829275).


CeaS1
CeaS2
S.co
S.av

Recently the crystal structure of CeaS2 complexed with TPP, Mg^{2+} and a sulfate anion was solved, and CeaS2 was found to be in a tetrameric form, composed of a dimer of two more tightly associated dimers (Caines *et al.*, 2004). In the same study, some sites that are thought to be involved in substrate binding and catalysis were also identified. The amino acid residues in CeaS2 that are proposed to be involved in the binding of TPP, G3P and L-arginine are (^{57}E, ^{413}F, ^{438}F, ^{463}D, ^{465}G ^{488}V, ^{490}N, ^{492}T, and ^{495}L), (^{36}R, ^{120}H, ^{121}Q, ^{271}Y, ^{410}I, ^{414}R, ^{415}H, ^{436}S, ^{495}L, and ^{496}I) and (^{113}D, ^{298}Y, ^{301}D and ^{303}R), respectively (Caines *et al.*, 2004). Similar binding sites for TPP (^{48}E, ^{395}F, ^{420}F, ^{445}D, ^{447}G ^{470}V, ^{472}N, and ^{477}L), G3P (^{27}R, ^{111}H, ^{112}Q, ^{259}Y, ^{392}I, ^{396}R, ^{397}H, ^{418}S, ^{477}L, and ^{478}I) and L-arginine (^{104}D, ^{286}Y, ^{289}D and ^{291}R) are also present in CeaS1, with the exception of ^{492}T from the TPP binding site in CeaS2, which is replaced by ^{474}R in CeaS1 (Figure 3.2.6).

The crystal structure of CeaS2 showed that the monomer consists of three domains (α, β and γ)(Caines *et al.*, 2004), which is very similar to the predicted tertiary structure of CeaS1 (Figure 3.2.7). Overall, the predicted tertiary structure of CeaS1 shows considerable similarity to the known crystal structure of the CeaS2 monomer (Figure 3.2.7). The main differences that can be noticed are the absence of the β-pleated sheet, β7 in CeaS1, and the elongation of the CeaS1 α-helix, α5 in comparison to CeaS2, as well as some minor differences in the lengths of a few loops found in the γ-domain of each protein (Figure 3.2.7). The high levels of similarities found in the protein sequences and structures of CeaS1 and CeaS2 suggests that CeaS1 probably forms a tetramer also, and functions in a manner similar to CeaS2, although these hypotheses need experimental verification.

3.3 Preparation and analysis of the *S. clavuligerus cea*S1 and *cea*S1/*cea*S2 mutants

In the previous section, the isolation and sequencing of a second copy of the gene encoding carboxyethylarginine synthase in *S. clavuligerus,* called *cea*S1, was described. The predicted CeaS1 is very similar to CeaS2, which is encoded in the clavulanic acid gene cluster. Due to this similarity, it was proposed that *cea*S1 might also be involved in clavulanic acid and/or 5*S* clavam metabolite biosynthesis. To test out this hypothesis, *cea*S1 single and *cea*S1/*cea*S2 double mutant strains of *S. clavuligerus* were prepared and

Figure 3.2.6 CeaS1 and CeaS2 protein alignment. Conserved, similar and dissimilar amino acid residues are shown on black, gray and white backgrounds, respectively. The symbols (•), (■) and (▼) indicate residues shown or proposed to be involved in thiamine pyrophosphate (TPP), glyceraldehyde-3-phosphate and L-arginine binding or catalysis respectively, in the crystal structure of CeaS2 (Caines *et al.* 2004), which are also found to be conserved in CeaS1. The arrow indicates the residue involved in TPP binding in CeaS2, which is different in the predicted CeaS1.

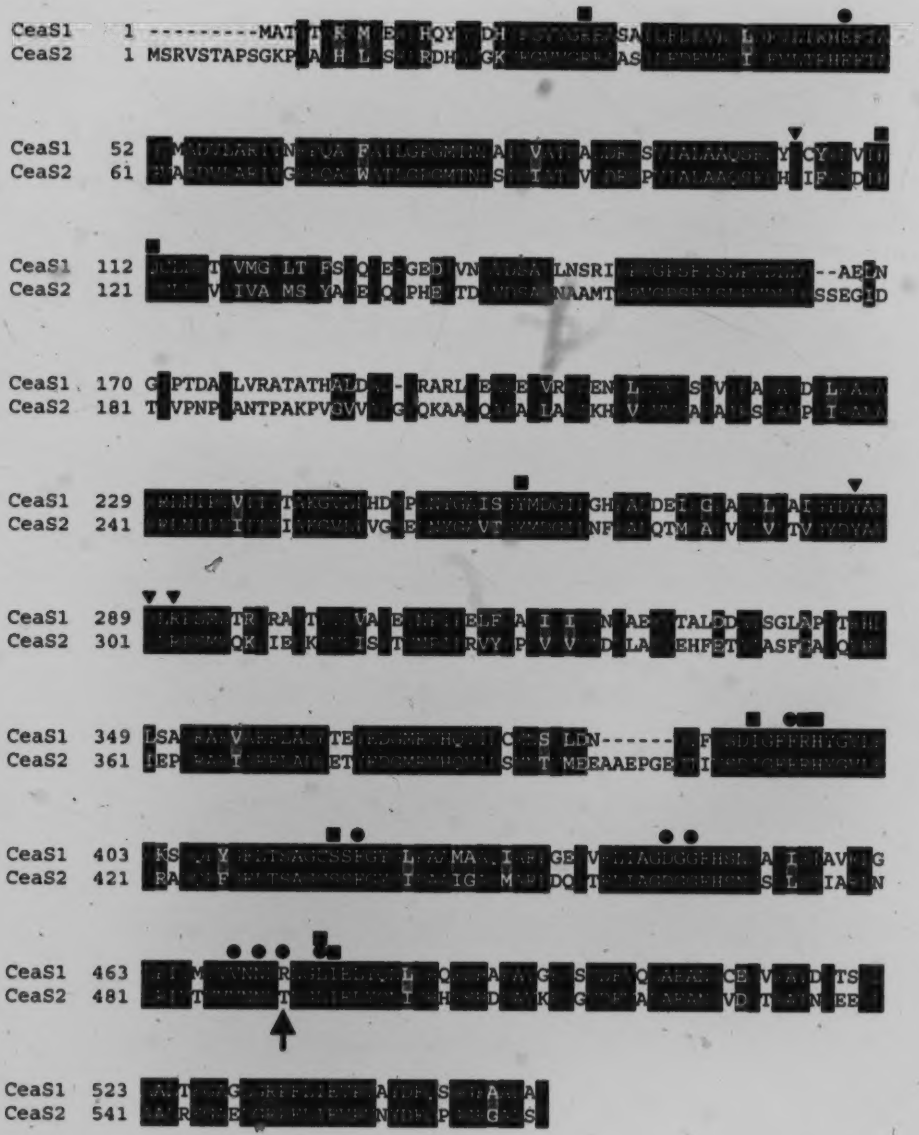
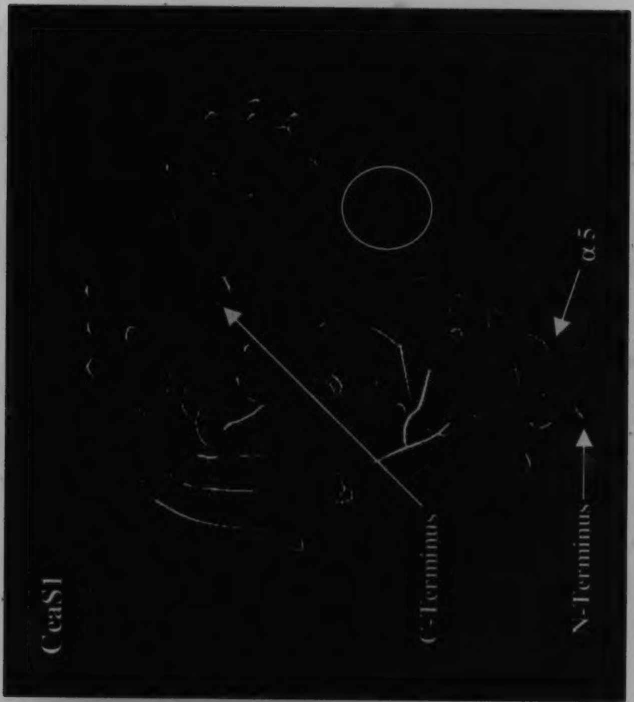
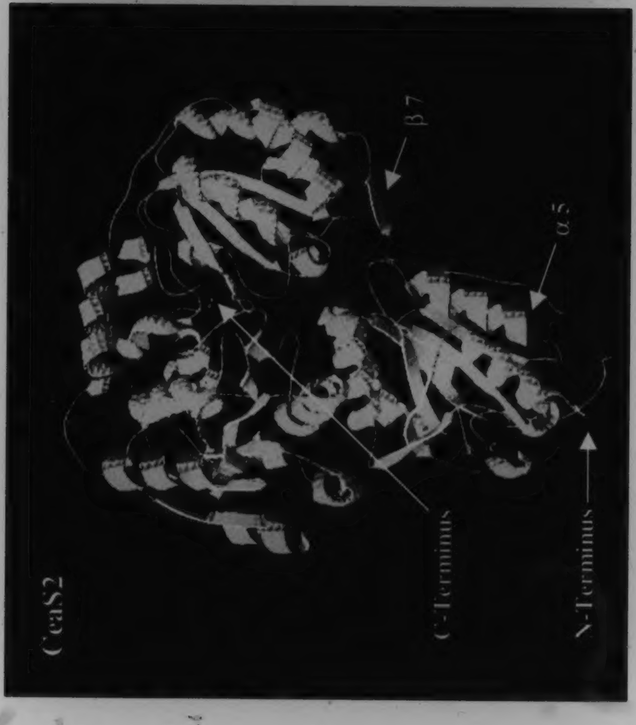

CeaS1 1
CeaS2 1
CeaS1 52
CeaS2 61
CeaS1 112
CeaS2 121
CeaS1 170
CeaS2 181
CeaS1 229
CeaS2 241
CeaS1 289
CeaS2 301
CeaS1 349
CeaS2 361
CeaS1 403
CeaS2 421
CeaS1 463
CeaS2 481
CeaS1 523
CeaS2 541

Figure 3.2.7 Predicted tertiary structure of CeaS1 and its comparison to CeaS2. The structure of CeaS1 was predicted by homology modeling. The crystal structure of CeaS2 has been solved (Caines *et al*. 2004), and the structure of the monomer was used for comparison. The amino (N) and carboxy (C) termini of the respective proteins are indicated. The different domains (α, β and γ) found in CeaS2 are also present in the predicted CeaS1, and are shown. Apparent differences in the structures such as the lack of β7 and extension of α5 in CeaS1 in comparison to CeaS2 are also indicated. Note that the discontinuity or break in the CeaS2 structure is due to unclear data obtained during crystallographic analysis (Caines *et al*. 2004).


CeaS1
C-Terminus
N-Terminus
α5


CeaS2
C-Terminus
N-Terminus
α5
β7

tested for their ability to produce clavulanic acid and the 5*S* clavams. Note: portions of this section have been published as part of the manuscript by Tahlan *et al.* (2004c).

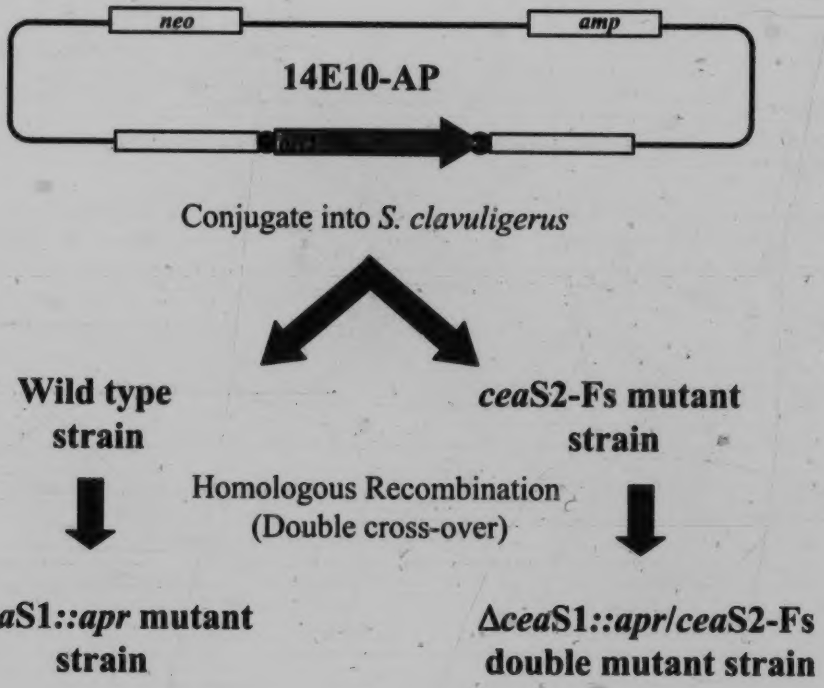
3.3.1 Preparation of the *cea*S1 mutant

The *S. clavuligerus cea*S1 mutant strain was prepared using the REDIRECT© PCR targeting system as described for the Δ*cea*S2*::apr* mutant (Section 3.1.2.1). The primers KTA14 and KTA15 were used to amplify the disruption cassette from the template plasmid pIJ773, and the cosmid 14E10 in the *E. coli* BW25113/pIJ780 strain was used to prepare the *cea*S1 disruption construct. The entire *cea*S1 gene was deleted from the cosmid 14E10 and replaced by the disruption cassette [*aac(3)IV+oriT*], to give the mutant cosmid 14E10-AP (Figure 3.3.1). The mutagenized cosmid 14E10-AP was then introduced into wild type *S. clavuligerus* by conjugation to isolate Δ*cea*S1*::apr* mutant strains as described earlier (Section 3.1.2.1).

Southern analysis was used to confirm that the chromosomal wild type copy of *cea*S1 had been replaced by the antibiotic resistance gene cassette. When a 777-bp *Eco*RI-*Nru*I fragment internal to *cea*S1 from p2.8-18, was used as the *cea*S1-specific probe, a 1.2 kb *Nco*I, and a 2.8 kb *Eco*RI fragment hybridized to the probe (Figure 3.3.2A) in the lanes containing either 14E10 cosmid DNA or wild type *S. clavuligerus* genomic DNA that had been digested with either *Nco*I or *Eco*RI. On the other hand, no fragments of these sizes hybridized to the *cea*S1-specific probe in lanes containing DNA from the Δ*cea*S1*::apr* mutant strains or 14E10-AP cosmid DNA. In addition, a 2.87 kb *Nco*I fragment also hybridized to the *cea*S1-specific probe in all the lanes that contained *S. clavuligerus* chromosomal DNA from either the wild type strain or the Δ*cea*S1*::apr* mutants (Figure 3.3.2A). These hybridizing bands can be attributed to the presence of the intact *cea*S2 gene in the respective strains, which cross-hybridizes to the *cea*S1-specific probe. The blot was then stripped and re-probed using the 1.38 kb *Eco*RI/*Hin*dIII fragment from pIJ773, which encodes the disruption cassette [*aac(3)IV+oriT*]. No hybridizing bands were observed in lanes containing 14E10 or wild type *S. clavuligerus* genomic DNA, whereas a 3.4 kb *Nco*I and a ~25 kb *Eco*RI fragment were seen to hybridize to the probe in lanes containing 14E10-AP or Δ*cea*S1*::apr* mutant

Figure 3.3.1 Preparation of the Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutants. (A) Strategy for the preparation of the Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutants using the cosmid 14E10-AP. The recombinant cosmid 14E10 was used to prepare the mutagenized cosmid 14E10-AP using the REDIRECT$^{\copyright}$ technology (see Figures 3.1.9 and 3.1.10). The cosmid 14E10-AP in which *cea*S1 was deleted and replaced by the apramycin resistance gene cassette [*acc(3)IV*] and conjugative origin of transfer (*oriT*), was then introduced into the *S. clavuligerus* wild type and the *cea*S2-Fs mutant strains to isolate the Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutants, respectively. The light gray bar represents the region of the *S. clavuligerus* chromosome flanking *cea*S1 (B) Differences in the restriction profiles in the region corresponding to the *S. clavuligerus* chromosome encompassing *cea*S1 in the wild type, Δ*cea*S1*::apr* and Δ*cea*S1*::apr*/*cea*S2-Fs mutant strains. Only restriction sites discussed in the text are shown. (A and B) The arrows represent the respective genes with the arrowheads representing the direction of transcription (see Figure 3.1.9).

A


neo
amp
14E10-AP
Conjugate into *S. clavuligerus*
Wild type strain
*cea*S2-Fs mutant strain
Homologous Recombination
(Double cross-over)
Δ*cea*S1::*apr* mutant strain
Δ*cea*S1::*apr*/*cea*S2-Fs double mutant strain

B

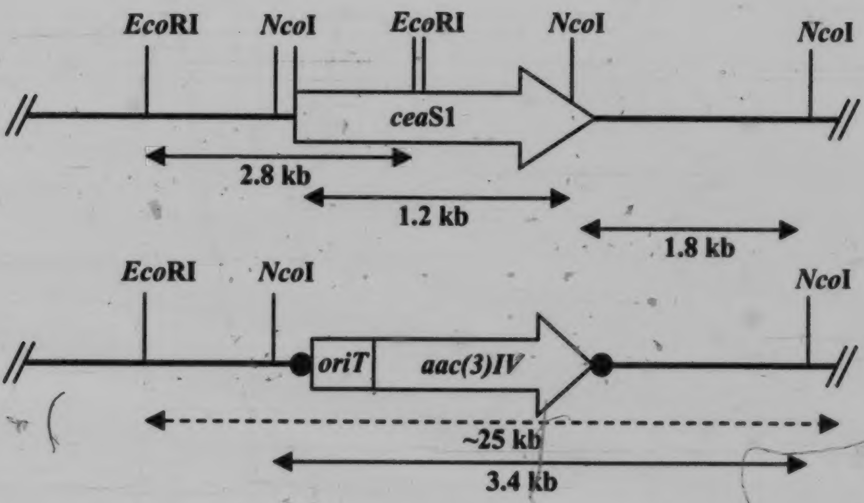
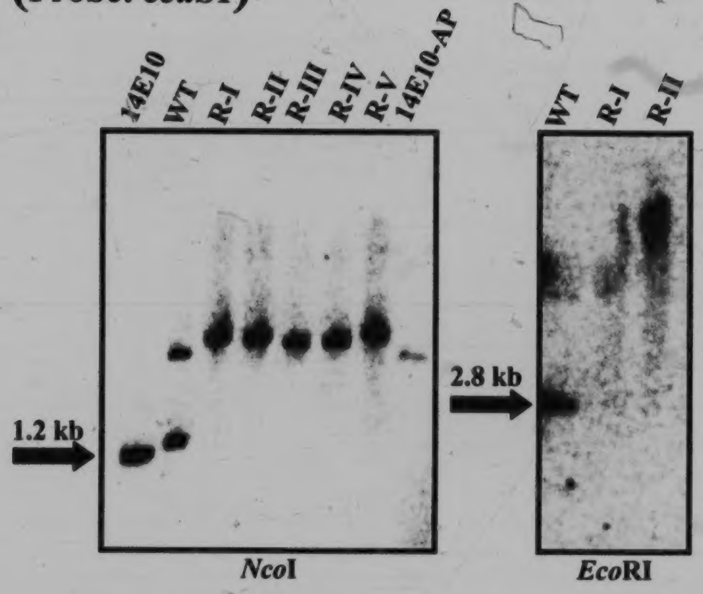
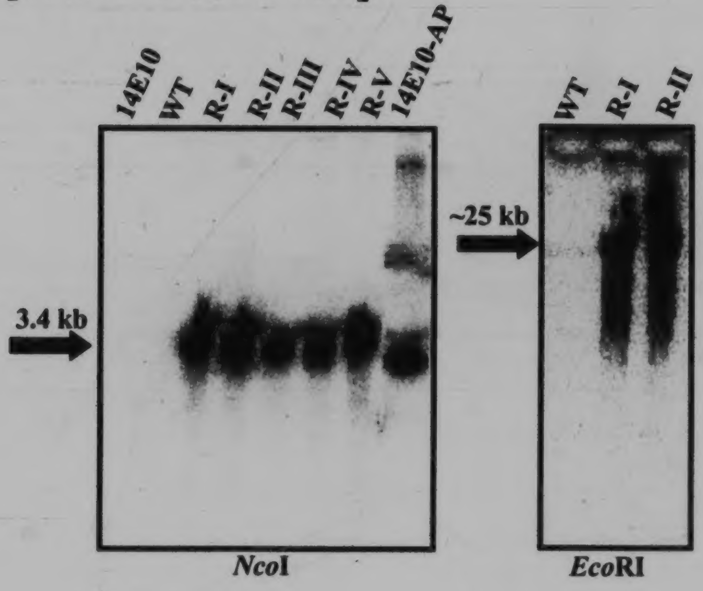

*Eco*RI
*Nco*I
*Eco*RI
*Nco*I
*Nco*I
*cea*S1
2.8 kb
1.2 kb
1.8 kb
*Eco*RI
*Nco*I
*Nco*I
oriT
aac(3)IV
~25 kb
3.4 kb

Figure 3.3.2 Verification of the Δ*cea*S1*::apr* single mutants by Southern analysis. Chromosomal DNA from the wild type strain and five Δ*cea*S1*::apr* mutants (R-I, R-II, R-III, R-IV and R-V) was digested with *Nco*I, or DNA from the wild type strain and two *cea*S1*::apr* mutants (R-I and R-II) was digested with *Eco*RI, before being subjected to Southern analysis. (A) The membrane was hybridized to a *cea*S1 specific probe. (B) A probe specific for the resistance cassette used to disrupt *cea*S1 in the mutants was used in the analysis. Recombinant cosmid DNA (14E10 and 14E10-AP) was included as controls. The sizes of the expected and hybridizing restriction fragments are indicated (see Figure 3.3.1). Some cross-hybridizing bands were also observed (A), which can be attributed to the presence of the *cea*S2 paralogue in the respective *S. clavuligerus* strains.



A (Probe: *cea*S1)

14E10
WT
R-I
R-II
R-III
R-IV
R-V
14E10-AP
1.2 kb
*Nco*I
WT
R-I
R-II
2.8 kb
*Eco*RI



B [Probe: *aac(3)IV+oriT*]

14E10
WT
R-I
R-II
R-III
R-IV
R-V
14E10-AP
3.4 kb
*Nco*I
WT
R-I
R-II
~25 kb
*Eco*RI

chromosomal DNA (Figure 3.3.2B). These hybridization profiles observed by Southern analysis were consistent with the replacement of the wild type chromosomal copy of *cea*S1 resulting in the Δ*cea*S1*::apr* mutant.

3.3.2 Phenotype of the Δ*cea*S1*::apr* mutant

Five Δ*cea*S1*::apr* mutants that were isolated and confirmed by Southern analysis (Figure 3.3.2) were also characterized by fermentation in SA and soy media. On HPLC analysis of SA and soy culture supernatants (Figure 3.3.3), a reduction in clavulanic acid production as compared to the wild type strain was observed (Tables 3.5), but it was not as severe as the reduction observed in the *cea*S2 mutants earlier (Section 3.1). Levels of 5*S* clavam production varied in soy medium and some mutants produced more 5*S* clavams than the wild type strain, but no specific pattern was identified (Table 3.6). The *cea*S1 disruption did not have any apparent effect on alanylclavam or cephamycin production as indicated by bioassays, and growth of all the strains in the fermentation media was comparable. The culture supernatants were also subjected to LC-MS analysis, which confirmed that all of the Δ*cea*S1*::apr* mutants still produced clavulanic acid in SA medium and both clavulanic acid and 5*S* clavams in soy medium (data not shown).

The Δ*cea*S1*::apr* (RI) mutant was also fermented in GSPG and in SA and soy media supplemented with glycerol or DHA as was done with the *cea*S2-Fs mutants (Section 3.1.1.2). The Δ*cea*S1*::apr* (RI) mutant produced only clavulanic acid in both glycerol-supplemented SA and soy, and in GSPG media (Table 3.7), whereas it produced the expected levels of the respective metabolites in DHA-supplemented SA and soy media (Table 3.8).

3.3.3 Preparation of the *cea*S1/*cea*S2 double mutant

Since both the *cea*S1 and the *cea*S2 single mutants produced at least some clavulanic acid and 5*S* clavams, a *cea*S1/*cea*S2 double mutant strain was prepared to verify that the two genes are indeed true paralogues. To prepare the *cea*S1/*cea*S2 double mutant strain, the *cea*S1 gene was knocked out in the previously prepared *cea*S2-Fs mutant strain (Section 3.1.1) using the REDIRECT© PCR targeting system as described in Section 3.2.1. The mutant cosmid 14E10-AP, containing Δ*cea*S1*::apr,* was introduced

Figure 3.3.3 Clavulanic acid and 5*S* clavam production by the *S. clavuligerus* Δ*cea*S1::*apr* mutants. Supernatants from the wild type and Δ*cea*S1::*apr* mutant cultures grown in SA and soy media for 96 hours were analyzed by HPLC. Of the five Δ*cea*S1::*apr* mutants fermented, chromatograms for only one (R-I) is shown as. Wild type soy culture supernatants were diluted (1/5) before analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.

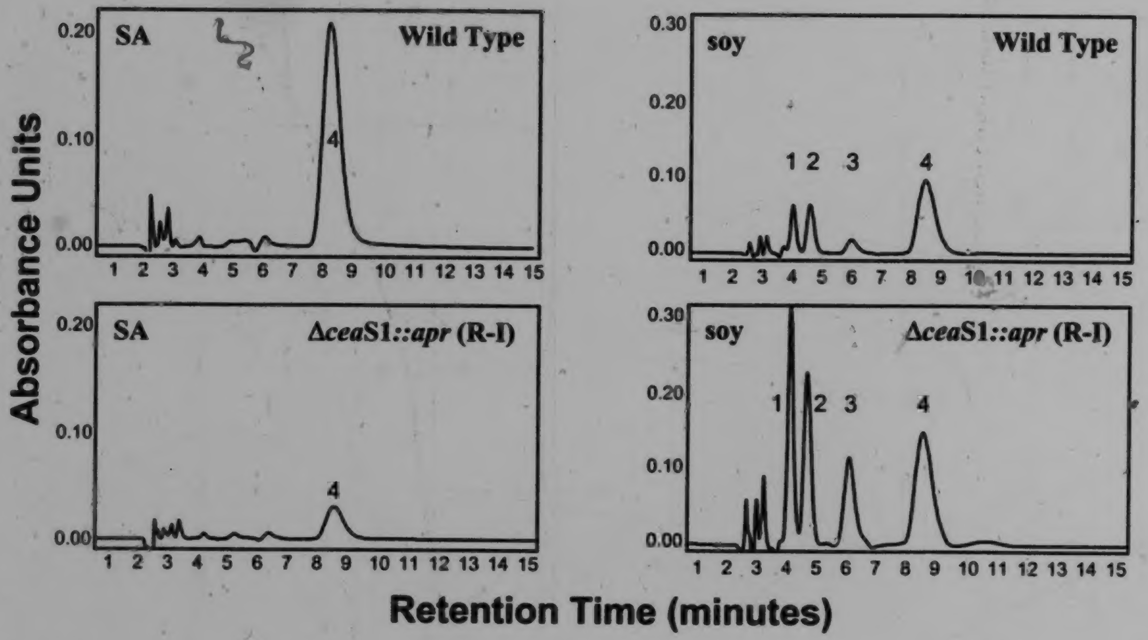

Absorbance Units
SA
Wild Type
0.20
0.10
0.00
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
soy
Wild Type
0.30
0.20
0.10
0.00
1 2 3 4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
SA
ΔceaS1::apr (R-I)
0.20
0.10
0.00
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
soy
ΔceaS1::apr (R-I)
0.30
0.20
0.10
0.00
1 2 3 4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Retention Time (minutes)

Table 3.5: Clavulanic acid produced by *S. clavuligerus cea*S1 mutants after 72 and 96 hours of growth on SA and soy media as determined by HPLC[a] analysis of culture supernatants.

Strain	Clavulanic acid (μg/ml)/(% Of wild type)[b]			
	SA medium		Soy medium	
	72 hours	96 hours	72 hours	96 hours
Wild type	120.5/100	82.3/100	181.5/100	187.2/100
Δ*cea*S1::*apr* (RI)	19.5/16.2	11.7/14.2	124.2/68.4	54.9/29.3
Δ*cea*S1::*apr* (RII)	34.6/28.7	12.6/15.3	128.8/70.9	128.0/68.4
Δ*cea*S1::*apr* (RIII)	23.6/19.6	13.7/16.6	73.4/40.4	104.9/56.0
Δ*cea*S1::*apr* (RIV)	16.2/13.4	26.7/32.4	104.9/57.8	2.4/1.3
Δ*cea*S1::*apr* (RV)	35.0/29.0	35.8/43.5	3.3/1.8	27.2/14.5

[a] Results from one fermentation experiment are shown and in some cases were also verified by LC-MS
[b] Clavulanic acid produced by the wild type strain was assigned a value of 100%

Table 3.6: Clavaminic acid, clavam-2-carboxylate and 2-hydroxymethyl clavam produced by *S. clavuligerus cea*S1 mutants after 72 and 96 hours of growth on soy medium as determined by HPLC[a] analysis of culture supernatants.

Strain	Clavaminic acid (µg/ml)/(% Of wild type)[b]		Clavam-2-carboxylate (µg/ml)/(% Of wild type)[b]		2-hydroxymethyl clavam (µg/ml)/(% Of wild type)[b]	
	72 hours	96 hours	72 hours	96 hours	72 hours	96 hours
Wild type	5.2/100	40.1/100	11.7/100	55.8/100	4.2/100	24.4/100
Δ*cea*S1::*apr* (RI)	5.9/113.4	38.5/96.0	18.4/157.3	35.6/63.8	6.2/147.6	27.4/112.3
Δ*cea*S1::*apr* (RII)	2.9/55.8	6.6/16.5	12.0/102.6	9.0/16.1	3.8/90.5	34.6/141.8
Δ*cea*S1::*apr* (RIII)	2.9/55.8	10.0/24.9	3.7/31.6	12.6/22.6	1.9/45.2	9.1/37.3
Δ*cea*S1::*apr* (RIV)	8.3/159.6	27.5/68.6	16.3/139.3	1.5/2.7	7.2/171.4	14.0/57.4
Δ*cea*S1::*apr* (RV)	22.5/432.6	27.9/69.6	ND[c]/0	26.2/47.0	6.9/164.3	22.0/90.2

[a] Results from one fermentation experiment are shown and in some cases were also verified by LC-MS

[b] Production of the respective metabolites by the wild type strain was assigned a value of 100%

[c] None detected

Table 3.7: Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*cea*S1::*apr* and Δ*cea*S1::*apr*/*cea*S2-Fs mutant strains grown in glycerol containing medium.

Strain	GSPG[a]		SA-G[b]		Soy-G[b]	
	5*S* clavams[c]	Clavulanic acid	5*S* clavams[c]	Clavulanic acid	5*S* clavams[c]	Clavulanic acid
Wild type[d]	+	+	-	+	+	+
Δ*cea*S1::*apr* (RI)[d]	-	+	-	+	-	+
Δ*cea*S1::*apr*/*cea*S2-Fs (2-3-R-1)[d]	-	-	-	-	-	-

[a] Supernatants from cultures grown on GSPG medium for 48 hours were analyzed by HPLC

[b] Supernatants from cultures grown in SA or soy medium for 96 hours, supplemented with 15 mg/ml glycerol (SA-G and Soy-G, respectively) were analyzed by bioassays

[c] The clavams are shown as a family of compounds as the bioassays only detected alanylclavam production

[d] Qualitative results are shown: +, the respective metabolites were detected and -, the respective metabolites were not detected

Table 3.8: Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*cea*S1::*apr* and Δ*cea*S1::*apr*/*cea*S2-Fs mutant strains as determined by HPLC analysis of culture supernatants, after 96 hours of growth in SA and soy media supplemented with dehydroacetic acid (DHA).

Strain	Clavaminic acid		Clavam-2-carboxylate (μg/ml)		2-hydroxymethyl clavam		Clavulanic acid	
	SA-D[a]	Soy-D[a]	SA-D[a]	Soy-D[a]	SA-D[a]	Soy-D[a]	SA-D[a]	Soy-D[a]
Wild type	0.17	18.7	1.52	50.42	1.86	92.25	1.73	117.45
Δ*cea*S1::*apr* (RI)	ND[b]	9.51	ND[b]	21.41	ND[b]	13.85	1.22	48.16
Δ*cea*S1::*apr*/*cea*S2-Fs (2-3-R-1)	ND[b]	ND[b]	ND[b]	ND[b]	ND[b]	ND[b]	ND[b]	ND[b]

[a] SA or soy medium supplemented with 0.5 mg/ml DHA (SA-D and Soy-D, respectively)

[b] None detected

into the *cea*S2-Fs mutant strain by conjugation (Figure 3.3.1), and homologous recombination between the mutagenized cosmid and the chromosome of the *S. clavuligerus* *cea*S2-Fs mutant strain led to the disruption of *cea*S1 in the desired background (Figure 3.3.1). Three parental *cea*S2-Fs (2-3, 2-7 and 2-8) mutants were used to isolate six *cea*S1/*cea*S2 mutants (referred to as the Δ*cea*S1*::apr/cea*S2-Fs mutants). Since *cea*S1 was knocked out in the *cea*S2-Fs background, the same DNA probes that were used to verify the Δ*cea*S1*::apr* mutant, were also used to confirm the disruption of *cea*S1 in the Δ*cea*S1*::apr/cea*S2-Fs mutants. The observed Southern hybridization profiles were the same as those expected for the Δ*cea*S1*::apr* mutant (Section 3.3.1), confirming that *cea*S1 gene disruption had taken place in the Δ*cea*S1*::apr/cea*S2-Fs double mutant (Figure 3.3.4).

3.3.4 Phenotype of the Δ*cea*S1*::apr/cea*S2-Fs double mutant

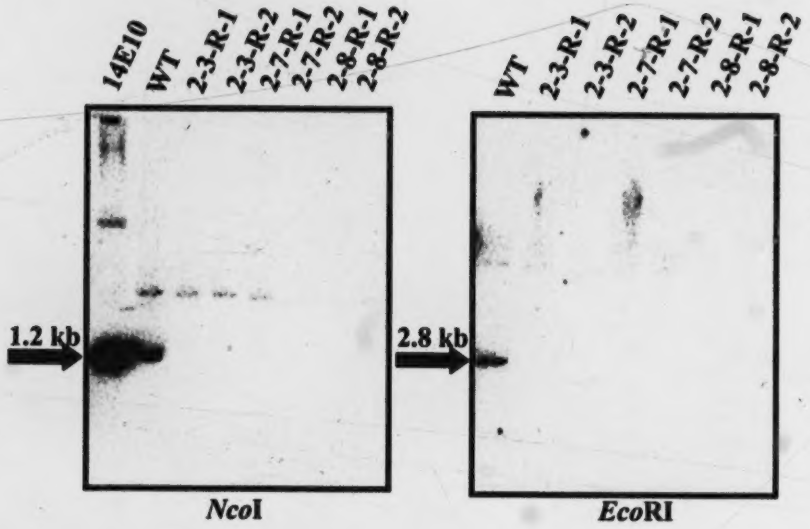
The six isolated Δ*cea*S1*::apr/cea*S2-Fs double mutants were analyzed for clavulanic acid, 5*S* clavam and cephamycin C production, after 72 and 96 hours of growth in SA and soy media. No clavulanic acid or 5*S* clavam production was detected in either SA or soy culture supernatants by HPLC (Figure 3.3.5) or by bioassays. In addition, the undulating baseline observed in the Δ*cea*S1*::apr/cea*S2-Fs mutant chromatograms in Figure 3.3.5, was due to non-specific background absorbance in the underivatized samples, and not due to clavam metabolite production by the mutants in soy medium. Bioassays also indicated that the Δ*cea*S1*::apr/cea*S2-Fs mutants were unaffected in cephamycin C biosynthesis.

The Δ*cea*S1*::apr/cea*S2-Fs (2-3-R-1) mutant was also fermented in SA and soy media supplemented with glycerol or DHA, and in GSPG medium as was done for the *cea*S2-Fs and Δ*cea*S1*::apr* mutants (Sections 3.1.1.2 and 3.3.2) . Under the conditions used, the Δ*cea*S1*::apr/cea*S2-Fs (2-3-R-1) mutant did not produce any detectable levels of either clavulanic acid or the 5*S* clavams in any of the media tested (Tables 3.7 and 3.8).

Since the parental *cea*S2-Fs mutants, which were used to prepare the Δ*cea*S1*::apr/cea*S2-Fs double mutants, typically produced clavulanic acid and 5*S* clavam

Figure 3.3.4 Southern analysis of Δ*cea*S1::*apr*/*cea*S2-Fs double mutant strains. Chromosomal DNA from the wild type strain and six Δ*cea*S1::*apr*/*cea*S2-Fs mutants (2-3-R-1, 2-3-R-2, 2-7-R-1, 2-7-R-2, 2-8-R-1 and 2-8-R-2) was digested with *Nco*I or *Eco*RI before being subjected to Southern analysis. (A) The membrane was hybridized to a *cea*S1 specific probe. (B) A probe specific for the resistance cassette used to disrupt *cea*S1 was used in the analysis. Recombinant cosmid DNA (14E10) was included as a control. The sizes of the expected and hybridizing restriction fragments are indicated (see Figure 3.3.1). (A) The cross hybridizing bands observed can be attributed to the presence of *cea*S2-Fs (carries a simple frame shift mutation in *cea*S2) in the respective *S. clavuligerus* strains.

A (Probe: *cea*S1)


14E10
WT
2-3-R-1
2-3-R-2
2-7-R-1
2-7-R-2
2-8-R-1
2-8-R-2
1.2 kb
NcoI
WT
2-3-R-1
2-3-R-2
2-7-R-1
2-7-R-2
2-8-R-1
2-8-R-2
2.8 kb
EcoRI

B [Probe: *aac(3)IV+oriT*]

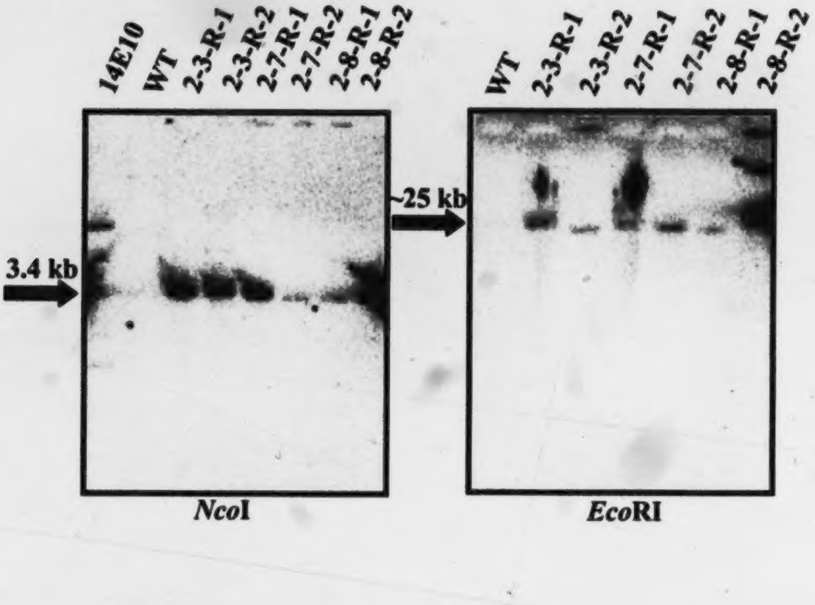
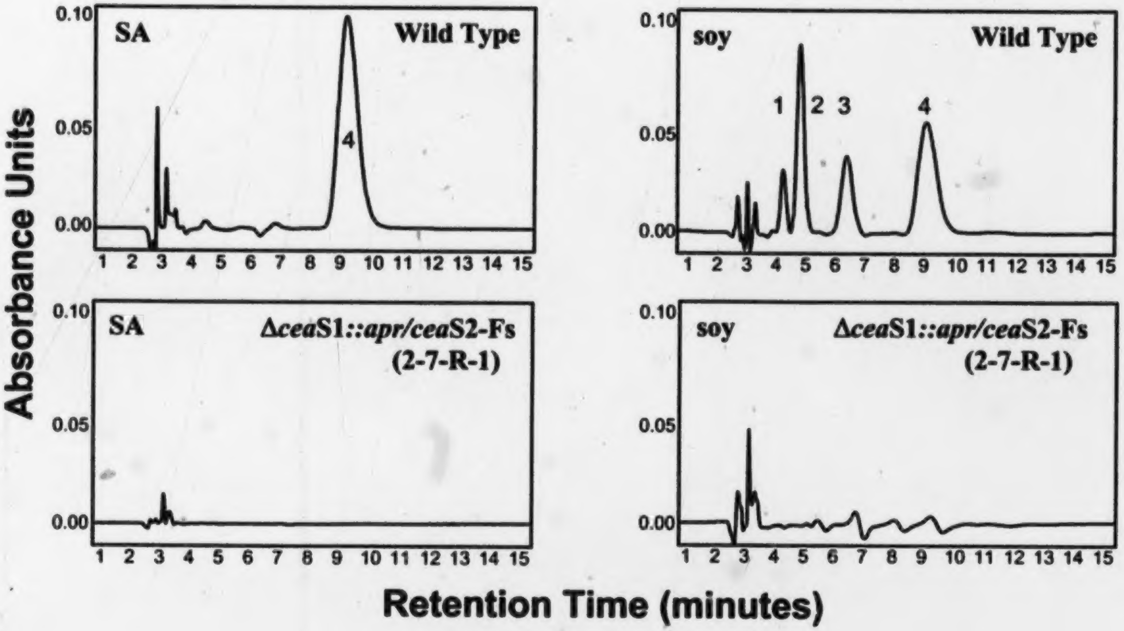

14E10
WT
2-3-R-1
2-3-R-2
2-7-R-1
2-7-R-2
2-8-R-1
2-8-R-2
3.4 kb
NcoI
WT
2-3-R-1
2-3-R-2
2-7-R-1
2-7-R-2
2-8-R-1
2-8-R-2
~25 kb
EcoRI

Figure 3.3.5 Clavulanic acid and 5*S* clavam production by the *S. clavuligerus* Δ*cea*S1*::apr/cea*S2-Fs mutants. Supernatants from the wild type and Δ*cea*S1*::apr/cea*S2-Fs mutant cultures grown in SA and soy media for 96 hours were analyzed by HPLC. Of the five Δ*cea*S1*::apr/cea*S2-Fs mutants fermented, chromatograms for only one representative strain (2-7-R-1) are shown. Wild type culture supernatants were diluted (1/5) before analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.


SA
Wild Type
4
soy
Wild Type
1
2
3
4
SA
ΔceaS1::apr/ceaS2-Fs
(2-7-R-1)
soy
ΔceaS1::apr/ceaS2-Fs
(2-7-R-1)
0.10
0.05
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Absorbance Units
Retention Time (minutes)

metabolites at such low levels that they were only detected by LC-MS analysis of soy culture supernatants (Section 3.1.1.2), the Δ*cea*S1*::apr/cea*S2-Fs mutants were also subjected to similar analysis, to determine if they still produced the respective metabolites. Soy culture supernatants from three Δ*cea*S1*::apr/cea*S2-Fs mutants (2-3-R1, 2-7-R1 and 2-8-R1) were analyzed by LC-MS, further confirming that the Δ*cea*S1*::apr/cea*S2-Fs mutants were completely blocked in clavulanic acid and 5*S* clavam metabolite biosynthesis, as none of the ions corresponding to the respective metabolites were detected (Figure 3.3.6). Note that the peak in chromatogram E (approximate retention time of 19.4 minutes and m/z =156) is not due to clavulanic acid as it lacked the correct absorbance spectrum and has a slightly shorter retention time (Figure 3.3.6).

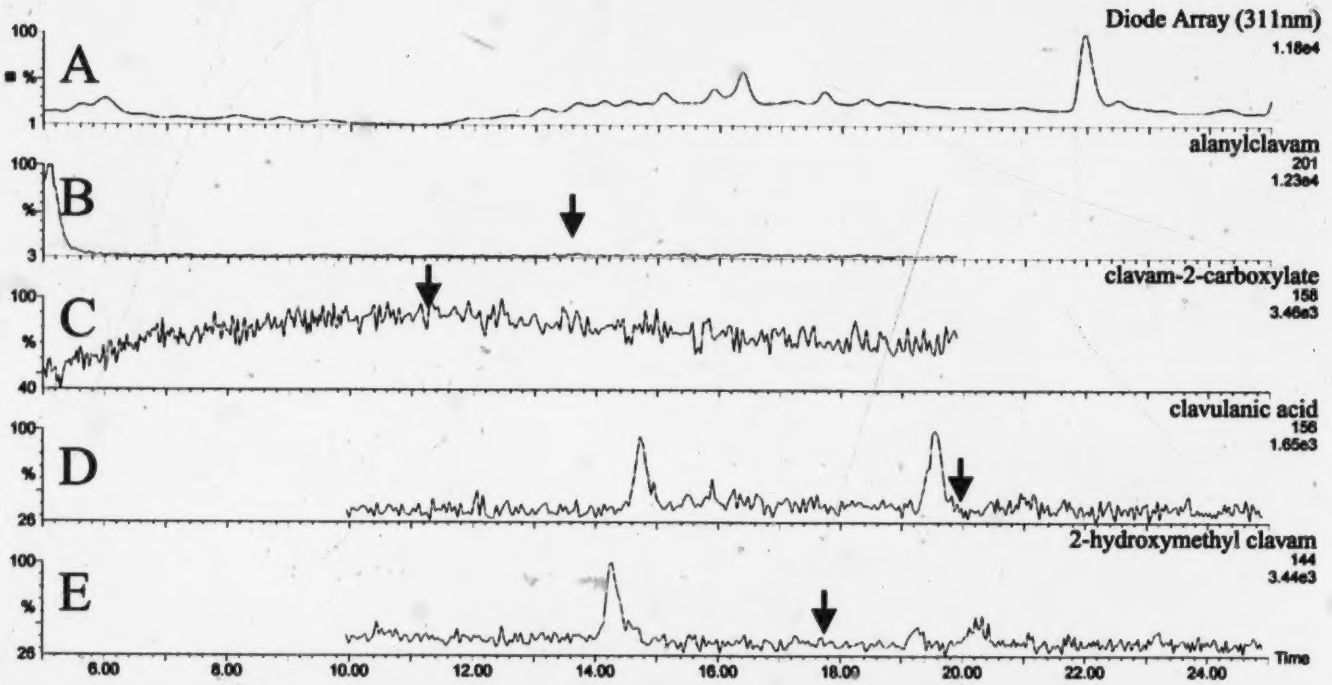
3.4 Preparation and analysis of the *c7p* and the *skn* mutants

When mutants defective in other genes from the paralogue gene cluster were prepared and analyzed, they had varying effects on clavulanic acid and 5*S* clavam metabolite biosynthesis. The *bls*1 (Tahlan *et al.*, 2004c) and *pah*1 (Jensen *et al.*, 2004b) genes have been shown to be involved in the biosynthesis of both clavulanic acid and 5*S* clavams, whereas the exact function of *oat*1 is still not known (Tahlan *et al.*, 2004c). Upstream of *oat*1, Jensen *et al.* (Unpublished) identified three additional ORFs (*c6p*, *c7p* and *skn*) (Figure 3.2.4), which encode putative proteins similar to aminotransferases, transcriptional regulators and sensor kinases from different sources, respectively. To determine if *c7p* or *skn* from the other end of the paralogue gene cluster were involved in clavulanic acid or 5*S* clavam metabolite biosynthesis, mutants defective in both genes were prepared and analyzed.

3.4.1 Sequence analysis of *c7p* and *skn*

The complete DNA sequence of both *c7p* and *skn* was obtained from the cosmid 14E10 by Jensen *et al.* (Unpublished), which also contains other genes from the paralogue gene cluster (Figure 3.2.4). Interestingly, *c7p* is very similar to the *cvm*7 gene (they share 50% and 33% end-to-end identity at the nucleotide and at the amino acid level, respectively), which also encodes a putative transcriptional regulator, and is located

Figure 3.3.6 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* Δ*cea*S1::*apr/cea*S2-Fs mutant (2-7-R-1). The chromatograms shown are as follows: (A) Absorbance at 311 nm due to clavam compounds derivatized with imidazole, (B) Profile due to derivatized alanylclavam fragmented ion, (C) Profile due to derivatized clavam-2-carboxylate fragmented ion, (D) Profile due to derivatized clavulanic acid fragmented ion and (E) Profile due to derivatized 2-hydroxymethyl clavam fragmented ion. The x-axis indicates the retention time with the arrows representing the missing peaks and the y-axis indicates the relative abundance of the respective peaks.


Diode Array (311nm)
1.18e4
A
100
%
1
alanylclavam
201
1.23e4
B
100
%
3
clavam-2-carboxylate
158
3.46e3
C
100
%
40
clavulanic acid
156
1.65e3
D
100
%
26
2-hydroxymethyl clavam
144
3.44e3
E
100
%
26
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time

in the clavam gene cluster of *S. clavuligerus* (Jensen *et al.*, Unpublished) (Figure 1.4.3). When a *cvm7* gene mutant was prepared and analyzed, it was found to have no apparent effect on clavulanic acid or 5*S* clavam production (Jensen *et al.*, Unpublished). The C7P and Cvm7 gene products are predicted to contain 818 and 1114 amino acids respectively, and both Cvm7 and C7P are similar to a family of novel transcriptional regulators, which includes PimR from *Streptomyces natalensis* (Anton *et al.*, 2004). The N-terminus of the predicted C7P protein contains SARP-like domains, which include BTAD (Bacterial Transcriptional Activator Domain) and the transcriptional regulatory protein C terminal domain (OmpR-type), which are also found in the DnrI/RedD/AfsR family of regulators (Section 1.5). In addition, the C-terminal region of the predicted C7P shows the presence of a weak AAA (ATPases Associated with diverse cellular Activities) type domain, which contains a type A, ATP/GTP binding motif.

The amino acid sequence of the predicted Skn was found to be similar to typical sensor kinase proteins from two component signal transduction systems from *S. coelicolor* and *S. avermitilis*. The C-terminal region of the predicted Skn contains a histidine kinase-like ATPase domain, whereas its N-terminus and central regions contain GAF (cGMP-specific and stimulated phosphodiesterases, *Anabaena* adenylate cyclases and *Escherichia coli* FhlA) domains, which are associated with the binding of small ligand molecules (Galperin, 2004). Due to the adjacent location of *c7p* and *skn*, mutants defective in both the genes were prepared and analyzed to determine if they were involved in regulating clavulanic acid and/or 5*S* clavam metabolite biosynthesis.

3.4.2 Preparation of the *c7p* and the *skn* mutants

The *S. clavuligerus c7p* and *skn* mutant strains were prepared using the previously described REDIRECT© PCR targeting system (Gust *et al.*, 2003). The primers pairs, cvm7par-For/cvm7par-Rev and sen-kin-For/sen-kin-Rev, along with the plasmid pIJ773 as template, were used in the preparation of the *c7p* and *skn* mutant strains, respectively, using the same strategy as described in the preparation of the Δ*cea*S2::*apr* mutants (Section 3.1.2.1). The mutant cosmids 14E10-C7P and 14E10-SKN were isolated where *c7p* and *skn* had been deleted, respectively, and replaced by the disruption cassette [*aac(3)IV+oriT*]. Both 14E10-C7P and 14E10-SKN were then conjugated into wild type *S. clavuligerus* where they gave rise to mutants defective in *c7p* and *skn* as described earlier (Section 3.1.2.1).

The replacement of the chromosomal wild type copies of *c7p* and *skn* by the disruption cassettes was confirmed by Southern hybridization. Genomic DNA from the wild type strain and six *c7p* and *skn* mutants was digested with *Nco*I and was used in the analysis. When a 1.7 kb *Sal*I fragment from p5N14E10Sal (internal to *c7p*), was used as the *c7p*-specific probe, a 5.24 kb *Nco*I fragment hybridized to the probe in lanes containing DNA from the wild type strain, where no hybridizing bands were observed in lanes containing DNA from the Δ*c7p::apr* mutant strains (Figure 3.4.1). The same blot was stripped and re-probed using the 1.4 kb *Eco*RI/*Hin*dIII fragment from pIJ773, encompassing the disruption cassette [*aac(3)IV+oriT*]. This time no hybridizing bands were observed in lanes containing DNA from the wild type strain, whereas a 4.18 kb fragment hybridized to the probe in lanes containing DNA from the Δ*c7p::apr* mutants (Figure 3.4.1).

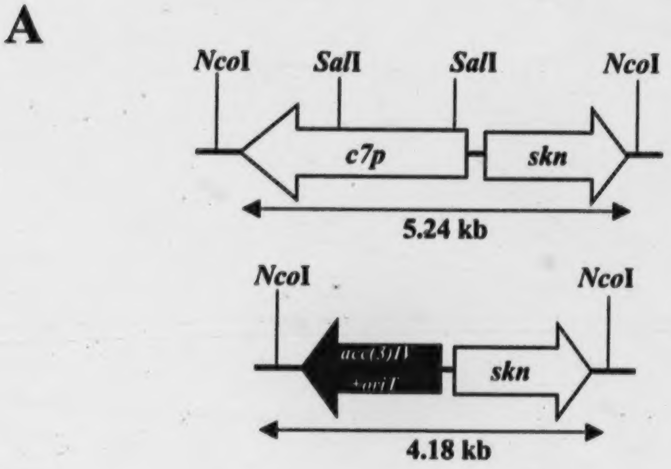
To verify the Δ*skn::apr* mutants, the 219 and 336 bp *Sal*I fragments from p5N14E10B, were used as the *skn*-specific probes. In the lane containing DNA from the wild type strain, a 5.2 kb *Nco*I fragment hybridized to the *skn*-specific probe, whereas no corresponding hybridizing bands were observed in lanes containing DNA from the Δ*skn::apr* mutants (Figure 3.4.2). When the same blot was stripped and re-probed using the *aac(3)IV+oriT*-specific probe, no hybridizing bands were observed in the wild type sample lane and a 4.97 kb *Nco*I fragment hybridized to the probe in lanes containing DNA from the Δ*skn::apr* mutants (Figure 3.4.2).

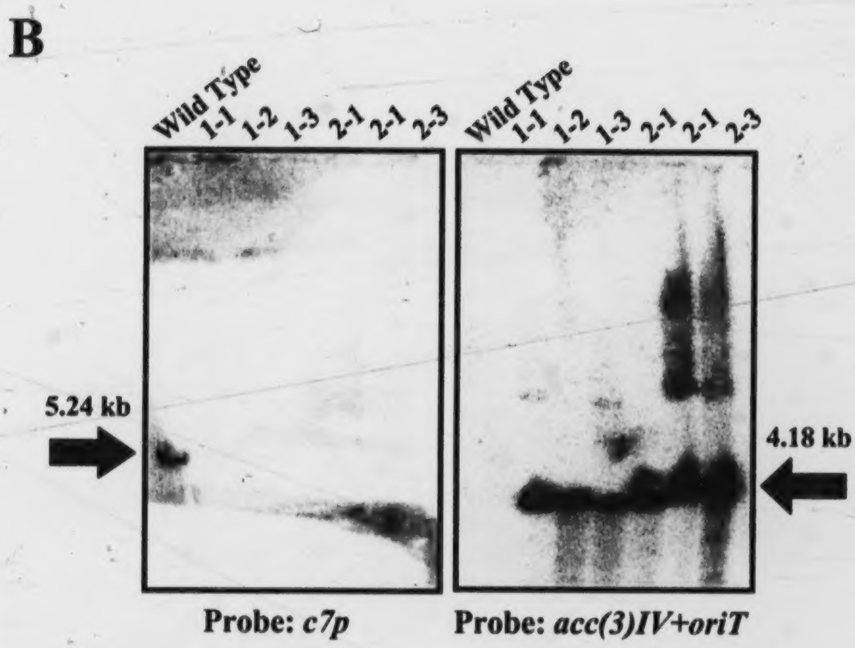
Overall, results from Southern analysis confirmed that the *c7p* and the *skn* genes had been disrupted by *aac(3)IV+oriT* in both the Δ*c7p::apr* and the Δ*skn::apr* mutants, respectively.

3.4.3 Phenotypes of the Δ*c7p::apr* and the Δ*skn::apr* mutants

Six Δ*c7p::apr* mutants were fermented in SA and soy media for 72 and 96 hours, respectively. HPLC analysis of soy culture supernatants showed that the Δ*c7p::apr* mutants did not produce any detectable levels of the 5*S* clavams, whereas they still

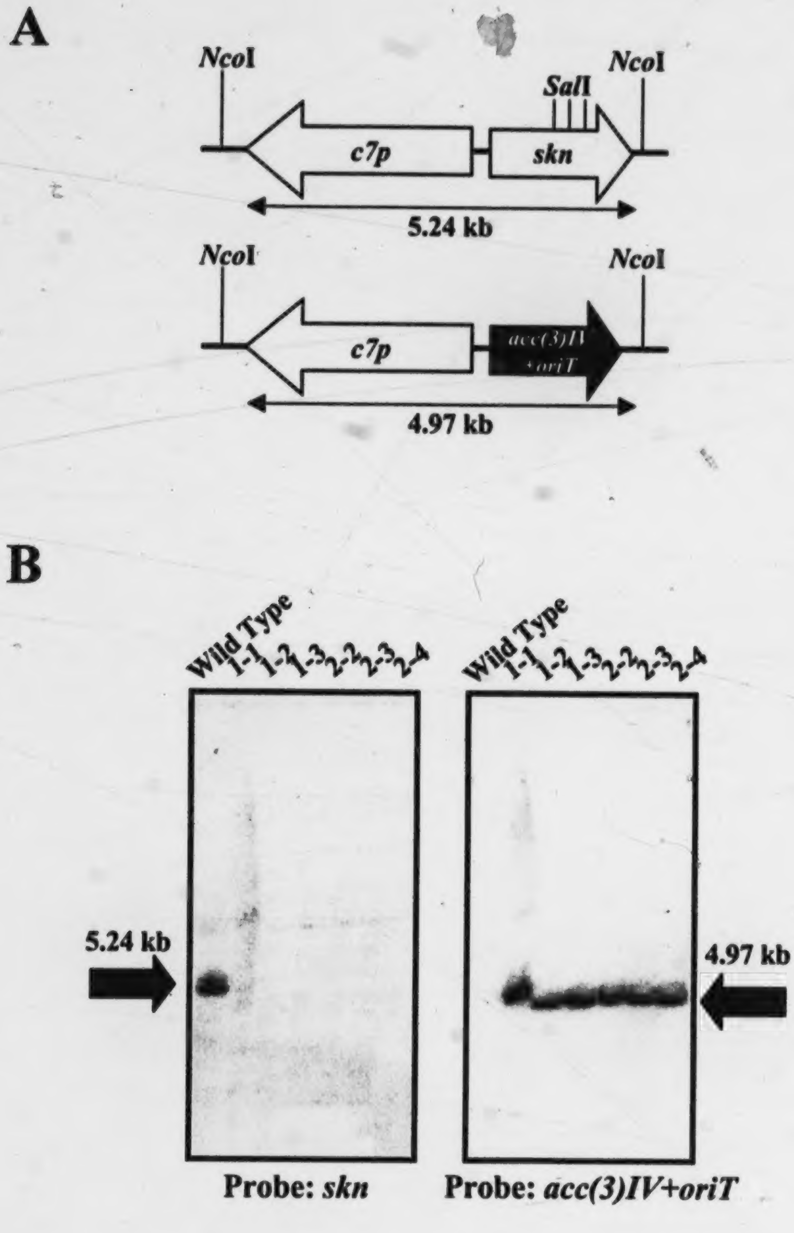
Figure 3.4.1 Southern analysis of the Δ*c7p::apr* mutants. (A) Diagrammatic representation of the *Nco*I fragment from the *c7p* region of the *S. clavuligerus* chromosome. The hollow arrows represent the respective genes with the arrowheads indicating the direction of transcription. The rest of the *S. clavuligerus* chromosome is represented by the thick line. The sizes of the corresponding restriction fragments (in kb), in the wild type and the Δ*c7p::apr* mutant strains are also indicated along with the restriction sites in *c7p* used for probe preparation. (B) DNA from wild type *S. clavuligerus* and six Δ*c7p::apr* mutants, 1-1, 1-2,1-3, 2-1, 2-2 and 2-3 was digested with *Nco*I before being subjected to Southern analysis. The membrane was probed using a probe specific for *c7p* followed by a probe specific for the disruption cassette used [*acc(3)IV+oriT*], after the membrane had been stripped of the first probe. The location of the hybridizing fragments matching the sizes of those shown in (A) are also indicated.


A
NcoI
SalI
SalI
NcoI
c7p
skn
5.24 kb
NcoI
NcoI
acc(3)IV +oriT
skn
4.18 kb


B
Wild Type
1-1
1-2
1-3
2-1
2-1
2-3
Wild Type
1-1
1-2
1-3
2-1
2-1
2-3
5.24 kb
4.18 kb

Probe: *c7p* **Probe:** *acc(3)IV+oriT*

Figure 3.4.2 Southern analysis of the *Δskn::apr* mutants. (A) Diagrammatic representation of the *Nco*I fragment from the *skn* region of the *S. clavuligerus* chromosome. The hollow arrows represent the respective genes with the arrowheads indicating the direction of transcription. The rest of the *S. clavuligerus* chromosome is represented by the thick line. The sizes of the corresponding restriction fragments in kb, in the wild type and the *Δskn::apr* mutant strains are also indicated and the restriction sites in *skn* used for probe preparation are also shown. (B) DNA from wild type *S. clavuligerus* and six *Δskn::apr* mutants, 1-1, 1-2,1-3, 2-2, 2-3 and 2-4 was digested with *Nco*I before being subjected to Southern analysis. The membrane was probed using a probe specific for *skn* followed by a probe specific for the disruption cassette [*acc(3)IV+oriT*], after the membrane had been stripped of the first probe. The location of the hybridizing fragments matching the sizes of those shown in (A) are also indicated.


A
NcoI
NcoI
SalI
c7p
skn
5.24 kb
NcoI
NcoI
c7p
acc(3)IV +oriT
4.97 kb
B
Wild Type
1-1
1-2
1-3
2-2
2-3
2-4
Wild Type
1-1
1-2
1-3
2-2
2-3
2-4
5.24 kb
4.97 kb
Probe: *skn*
Probe: *acc(3)IV+oriT*

produced clavulanic acid (Figure 3.4.3). Under the same conditions, the wild type strain produced both clavulanic acid and the 5*S* clavams in soy medium (Figure 3.4.3). In SA medium, both the wild type and the Δ*c7p::apr* mutants produced clavulanic acid (Table 3.9). Overall, the Δ*c7p::apr* mutants and the wild type strains produced similar levels of clavulanic acid in both SA and soy media (Table 3.9). Bioassays also showed that the Δ*c7p::apr* mutants did not produce any detectable alanylclavam in soy medium, whereas the production of clavulanic acid and cephamycin C was not affected in the same mutants.

Supernatants from 96 hour Δ*c7p::apr* (1-1 and 2-1) mutant soy cultures were also analyzed by LC-MS (Figure 3.4.4). Ions corresponding to the 5*S* clavams were not detected in the culture supernatants except for the 1-1 strain, which showed the presence of a minor 2-hydroxymethyl clavam peak (data not shown). The same culture supernatants showed a clear peak corresponding to clavulanic acid (Figure 3.4.4).

Six Δ*skn::apr* mutants were also fermented as described above. HPLC analysis revealed that none of the mutants produced any detectable levels of the 5*S* clavams in soy medium in comparison to the wild type strain (Figure 3.4.3). In addition, the Δ*skn::apr* mutants were not affected in clavulanic acid biosynthesis, as it was detected in both soy and SA culture supernatants from the Δ*skn::apr* mutants and the wild type strain (Table 3.10). Overall, the Δ*skn::apr* mutants produced slightly lower or similar levels of clavulanic acid in comparison to the wild type strain (Table 3.10). Bioassays reconfirmed results obtained by HPLC and showed that the Δ*skn::apr* mutants did not produce any detectable alanylclavam, but still produced cephamycin C and clavulanic acid at levels comparable to the wild type strain in both soy and SA media.

Supernatants from 96-hour soy cultures from the Δ*skn::apr* (1-1 and 2-1) mutants were also analyzed by LC-MS. Results clearly indicated that the mutants did not produce any detectable 5*S* clavams but still retained the ability to produce clavulanic acid (Figure 3.4.5).

3.5 Examination of the linkage between the clavulanic acid, clavam and paralogue gene cluster

Figure 3.4.3 Clavulanic acid and 5*S* clavams produced by the *S. clavuligerus* wild type, Δ*c7p::apr* and Δ*skn::apr* mutant strains as determined by HPLC analysis. Soy culture supernatants from the respective strains were analyzed after 96 hours of growth as described earlier. Out of each of the six Δ*c7p::apr* and Δ*skn::apr* mutants fermented, chromatograms for only one mutant defective in each gene are shown. All culture supernatants were diluted before HPLC analysis. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid. Note that results obtained by HPLC were also verified by LC-MS analysis of 96-hour soy culture supernatants.

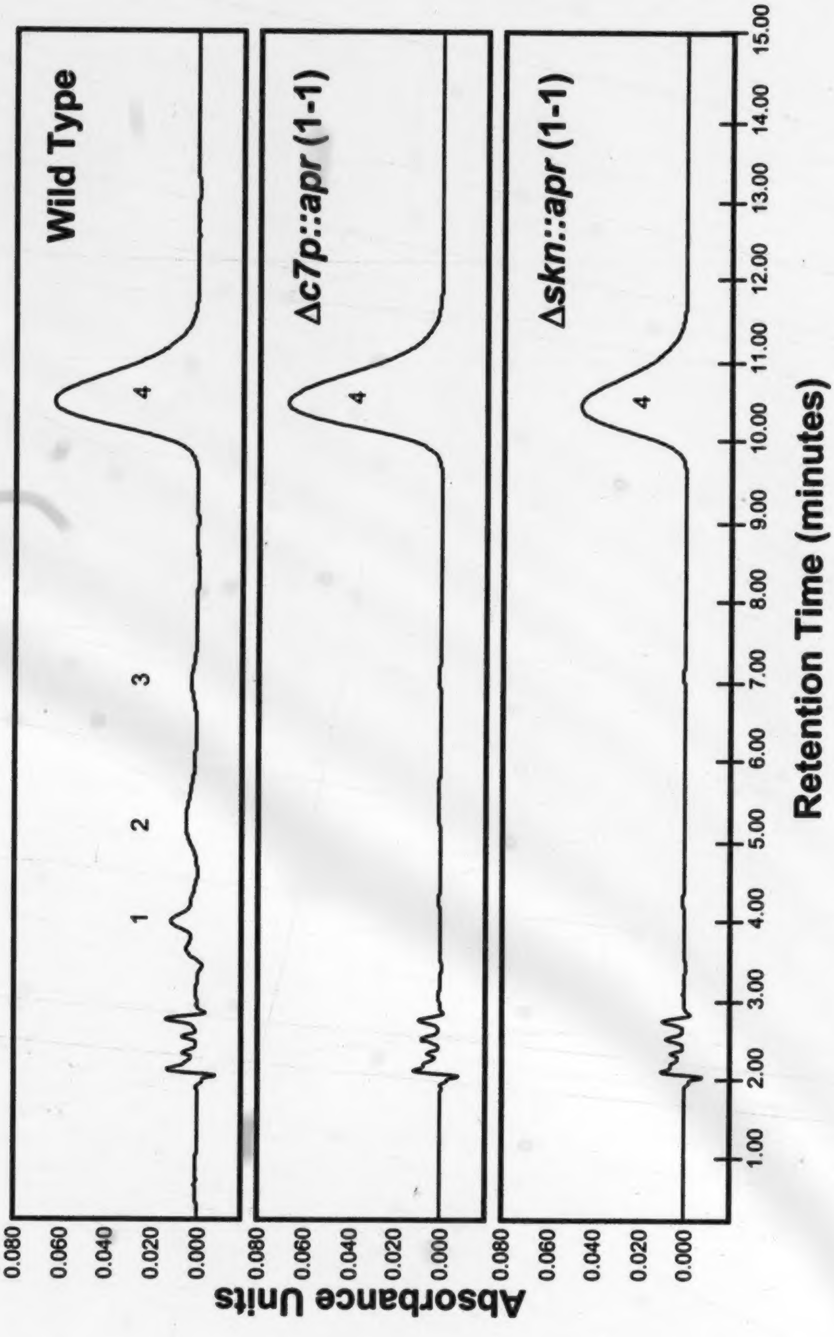

Wild Type
Δc7p::apr (1-1)
Δskn::apr (1-1)
1
2
3
4
4
4
0.080
0.060
0.040
0.020
0.000
Absorbance Units
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
Retention Time (minutes)

Table 3.9: Clavulanic acid produced by *S. clavuligerus* Δ*c7p::apr* mutants after 72 and 96 hours of growth on SA and soy medium as determined by HPLC[a] analysis of culture supernatants.

Strain	Clavulanic acid (μg/ml)/(% Of wild type)[b]			
	SA medium		Soy medium	
	72 hours	96 hours	72 hours	96 hours
Wild type	9.8/100	14.6/100	111/100	143/100
Δ*c7p::apr* (1-1)	11.6/118.4	9.7/66.4	105/94.6	151/105.6
Δ*c7p::apr* (1-2)	10.8/110.2	7.5/51.4	103/92.8	120/83.9
Δ*c7p::apr* (1-3)	9.2/93.9	7.3/50.0	110/99.1	151/105.6
Δ*c7p::apr* (2-1)	5.9/60.2	4.4/30.1	54/48.6	115/80.4
Δ*c7p::apr* (2-2)	6.8/69.4	6.9/47.3	92/82.9	208/145.5
Δ*c7p::apr* (2-3)	8.4/85.7	1.7/11.7	129/116.2	163/114.0

[a] Results from one fermentation experiment are shown and in some cases were also verified by LC-MS

[b] Clavulanic acid produced by the wild type strain was assigned a value of 100%

Figure 3.4.4 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* *Δc7p::apr* mutant (2-1). The chromatograms shown are as follows: (A) Absorbance at 311 nm due to clavam compounds derivatized with imidazole, (B) Profile due to derivatized alanylclavam fragmented ion, (C) Profile due to derivatized clavam-2-carboxylate fragmented ion, (D) Peak due to derivatized clavulanic acid fragmented ion and (E) Profile due to derivatized 2-hydroxymethyl clavam fragmented ion. The x-axis indicates the retention time with the arrow indicating the clavulanic acid peak observed, and the y-axis indicates the relative abundance of the respective peaks.

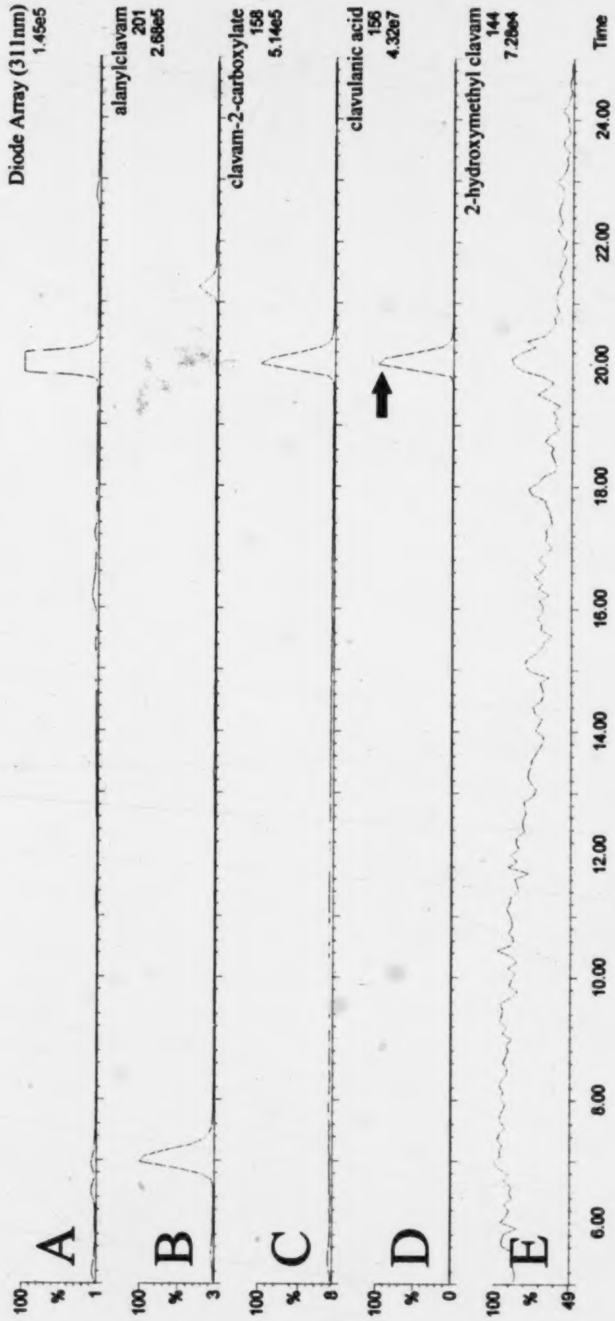

Diode Array (311nm)
1.45e5
alanylclavam
201
2.68e5
clavam-2-carboxylate
158
5.14e5
clavulanic acid
156
4.32e7
2-hydroxymethyl clavam
144
7.28e4
A
B
C
D
E
100
%
1
100
%
3
100
%
8
100
%
0
100
%
49
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time

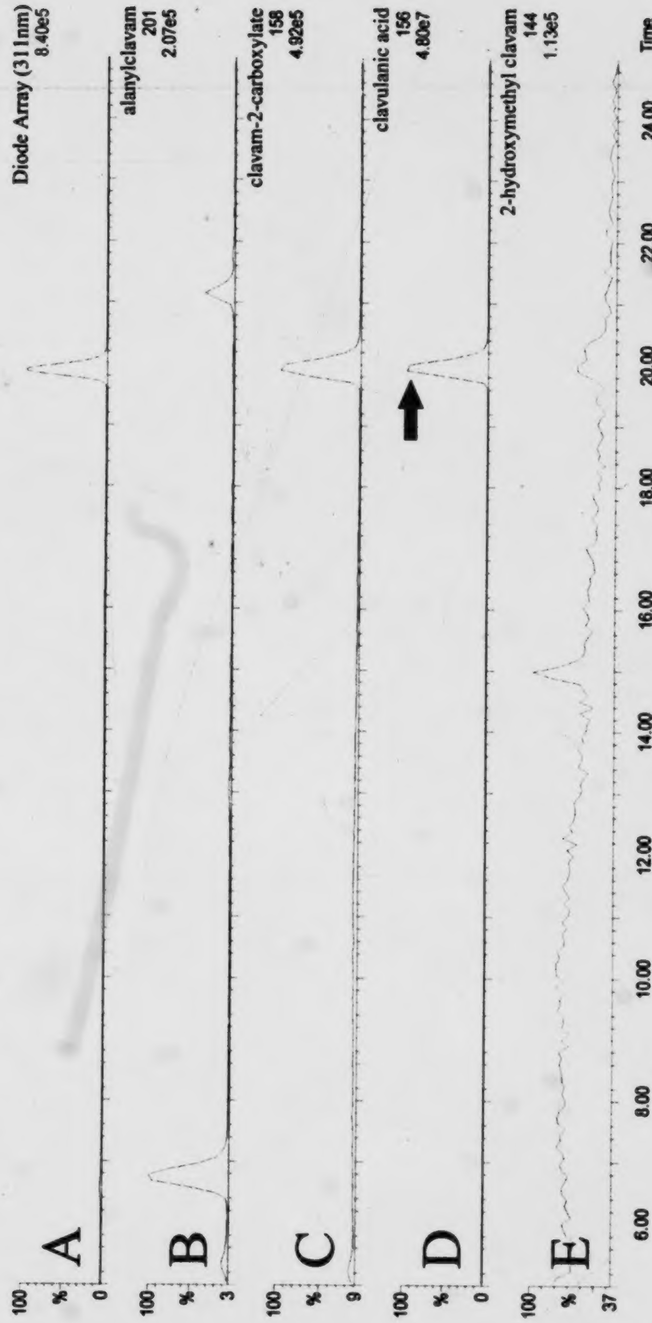
Table 3.10: Clavulanic acid produced by *S. clavuligerus* Δ*skn::apr* mutants after 72 and 96 hours of growth on SA and soy medium as determined by HPLC[a] analysis of culture supernatants.

Strain	Clavulanic acid (μg/ml)/(% Of wild type)[b]			
	SA		soy	
	72 hours	96 hours	72 hours	96 hours
Wild type	9.8/100	14.6/100	111/100	143/100
Δ*skn::apr* (1-1)	11.3/115.3	4.3/29.5	88/79.3	102/71.3
Δ*skn::apr* (1-2)	20.3/207.2	4.0/27.4	116/104.5	110/76.9
Δ*skn::apr* (1-3)	10.1/103.1	3.6/24.7	91/82.0	86/60.1
Δ*skn::apr* (2-2)	8.1/82.7	3.9/26.7	73/65.8	130/90.9
Δ*skn::apr* (2-3)	9.5/96.9	11.6/79.5	133/119.8	130/90.9
Δ*skn::apr* (2-4)	11.4/116.3	3.7/25.3	104/93.7	99/69.2

[a] Results from one fermentation experiment are shown and in some cases were also verified by LC-MS

[b] Clavulanic acid produced by the wild type strain was assigned a value of 100%

Figure 3.4.5 LC-MS analysis of soy culture supernatants from the *S. clavuligerus* Δ*skn::apr* mutant (2-1). The chromatograms shown are as follows: (A) Absorbance at 311 nm due to clavam compounds derivatized with imidazole, (B) Profile due to derivatized alanylclavam fragmented ion, (C) Profile due to derivatized clavam-2-carboxylate fragmented ion, (D) Peak due to derivatized clavulanic acid fragmented ion and (E) Profile due to derivatized 2-hydroxymethyl clavam fragmented ion. The x-axis indicates the retention time with the arrow indicating the clavulanic acid peak observed, and the y-axis indicates the relative abundance of the respective peaks.


Diode Array (311nm)
8.40e5
alanylclavam
201
2.07e5
clavam-2-carboxylate
158
4.92e5
clavulanic acid
156
4.80e7
2-hydroxymethyl clavam
144
1.13e5
A
B
C
D
E
100
%
0
100
%
3
100
%
9
100
%
0
100
%
37
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time

Clavulanic acid and the 5*S* clavams are known to share a common biosynthetic pathway up to the level of clavaminic acid (Egan *et al.*, 1997)(Figure 1.4.1), but at least three groups of genes have been shown to be involved in their biosynthesis (Figure 3.5.1). The clavulanic acid arm of the ceph-clav supercluster (Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Li *et al.*, 2000; Mellado *et al.*, 2002)(Figure 1.4.2), the clavam gene cluster (Mosher *et al.*, 1999)(Figure 1.4.3) and the paralogue gene cluster (Jensen *et al.*, 2004b; Tahlan *et al.*, 2004c)(Figure 3.2.4), all contain some genes that are involved in the biosynthesis of both clavulanic acid and the 5*S* clavams.

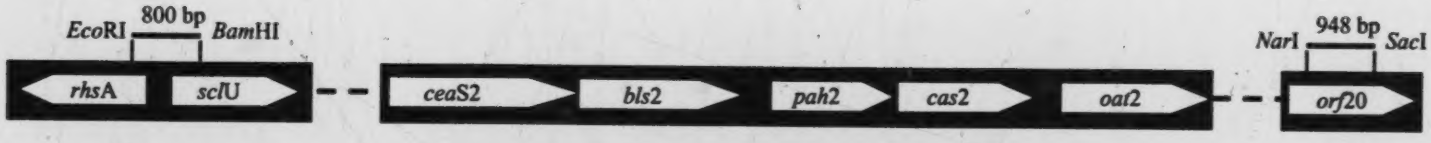
Since genes for antibiotic biosynthesis are invariably clustered in bacterial antibiotic producers, we examined the linkage between the ceph-clav, the clavam and the paralogue gene clusters to investigate their relative locations on the *S. clavuligerus* chromosome. Chromosome walking and PFGE (Pulsed Field Gel Electrophoresis) studies were carried out to test the hypothesis that the three gene clusters would be linked to form a still larger supercluster. In addition, results from linkage studies are expected to provide further insight into the relationship between the respective gene clusters and the biosynthetic pathways leading to clavulanic acid and the 5*S* clavams. Note: portions of this section have been published as part of the manuscript by Tahlan *et al.* (2004b).

3.5.1 Chromosome walking analysis

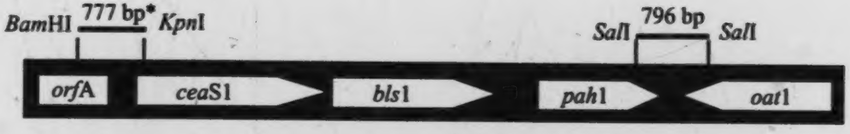
Previous studies involving Southern analysis of wild type *S. clavuligerus* chromosomal DNA fractionated by PFGE suggested that the clavulanic acid and the paralogue gene clusters might be physically linked, as a 180 kb *Ase*I macro restriction fragment hybridized to probes specific for both gene clusters (H-U. Park, Personal communication). With this in mind, we used chromosome walking to see if cosmids containing genes from the clavulanic acid, the paralogue and the clavam clusters, could be linked physically.

Before commencing chromosome-walking studies, we needed to isolate cosmids encompassing the clavulanic acid, the paralogue and the clavam gene clusters, respectively (Figure 3.5.1). A pWE15-based cosmid library containing *S. clavuligerus* chromosomal DNA inserts of approximately 42 kb was used for this purpose (Jin *et al.*, 2004). Recombinant cosmids encompassing the clavulanic acid gene cluster were isolated

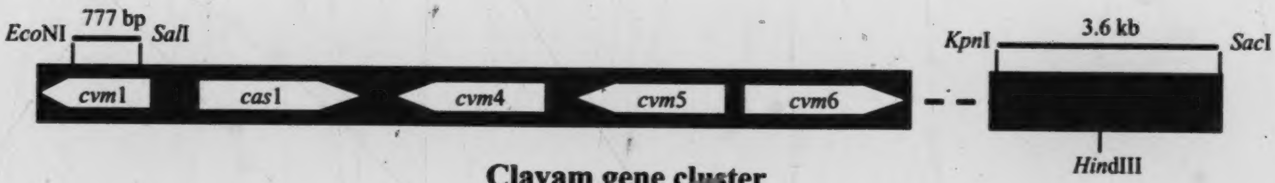
Figure 3.5.1 Diagrammatic representation of the cephamycin-clavulanic acid, the paralogue and the clavam gene clusters. The solid box represents the *S. clavuligerus* chromosome and the hollow arrows represent the genes with the arrowheads indicating the direction of transcription. Restriction sites in the cephamycin-clavulanic, paralogue and clavam gene clusters used in isolating probes for Southern analysis are also shown. The 777 bp *Bam*HI/*Kpn*I DNA fragment used as the paralogue gene cluster specific probe is also shown, and (*) indicates that the restriction sites used to isolate the probe were engineered, and are not present in the paralogue gene cluster. The *Hin*dIII restriction site present downstream of the clavam gene cluster is also shown. Note that the entire cephamycin-clavulanic acid and clavam gene clusters are not shown for the sake of simplicity, and the missing portions of the respective gene clusters are indicated by the dashed lines (Diagram is not to scale).


EcoRI
800 bp
BamHI
rhsA
sclU
ceaS2
bls2
pah2
cas2
oat2
NarI
948 bp
SacI
orf20

Cephamycin-Clavulanic acid gene clusters


BamHI
777 bp*
KpnI
SalI
796 bp
SalI
orfA
ceaS1
bls1
pah1
oat1

Paralogue gene cluster


EcoNI
777 bp
SalI
cvm1
cas1
cvm4
cvm5
cvm6
KpnI
3.6 kb
SacI
HindIII

Clavam gene cluster

using an *orf*20-specific probe (Figure 3.5.2), consisting of a 948 bp *Nar*I/*Sac*I DNA fragment internal to *orf*20 from pEB3 (Figure 3.5.1, Table 2.3). It was already known that the recombinant cosmids 6G9 and 14E10 from the same library (Table 2.3) contained *pah*1 (Jensen *et al*., 2004b), in addition to other genes from the paralogue gene cluster (Tahlan *et al*., 2004c)(Figure 3.5.2, Section 3.2.1). To isolate additional cosmids containing DNA from the paralogue gene cluster, the cosmid library was screened using the 777 bp *Bam*HI/*Kpn*I DNA fragment from pTOPO-*cea*S1-4 as the probe (Table 2.3), which predominantly encompasses the region immediately upstream of *cea*S1 from the paralogue gene cluster (Figure 3.5.1). This led to the isolation of the recombinant cosmid 4D3 in addition to 6G9 and 14E10, all of which hybridized to the paralogue gene cluster-specific probe (Figure 3.5.2, Table 3.11). The library was also screened for recombinant cosmids containing DNA from the clavam gene cluster using the 777 bp *Eco*NI/*Sal*I fragment from pCEC026, which encodes *cvm*1 (Figure 3.5.1), as the probe for the clavam gene cluster (Table 2.3). None of the cosmid clones present in the library hybridized to the clavam gene cluster-specific probe, and the hybridizations were repeated once more, to ensure that none of the cosmids representing the clavam gene cluster were missed. Again, no hybridization was observed, indicating the region of the *S. clavuligerus* chromosome encompassing the clavam gene cluster was not represented, in the fraction of the cosmid library subjected to analysis.

Chromosome walking was then used to identify physically linked, contiguous series of cosmids containing *S. clavuligerus* chromosomal DNA from the region downstream of clavulanic acid gene cluster, as it was already known that the cephamycin C gene cluster was situated upstream of the clavulanic acid gene cluster in the ceph-clav supercluster (Aidoo *et al*., 1994; Hodgson *et al*., 1995; Ward and Hodgson, 1993)(Figure 3.5.1). Since *orf*20 represents the end-most gene in the clavulanic acid part of the sequenced ceph-clav supercluster (Jensen *et al*., 2004a)(Figure 3.5.1), four subsequent chromosome walks were conducted to isolate cosmids extending beyond this region (Table 3.11). The cosmid 12E10, which contains *orf*20, and additional cosmid clones identified in the subsequent walks, were used to prepare probes for the next walk by digesting the clones with *Sac*II, which cuts frequently within *S. clavuligerus* genomic DNA but not within the cosmid vector. Self-ligation resulted in removal of most of the

Figure 3.5.2 Cosmids containing the cephamycin-clavulanic acid and paralogue gene clusters. (A) Diagrammatic representation of the relative locations and orientations of the cosmids isolated during chromosome walking in the region flanking the cephamycin-clavulanic acid (ceph-clav) gene cluster. The white arrows represent the *rhs*A gene from the cephamycin end and the *orf*20 gene from the clavulanic acid end of the ceph-clav gene cluster. (B) Cosmids present in the library that contain the paralogue gene cluster. The white arrows represent the *cea*S1 and the *skn* genes from the two ends of the paralogue gene cluster. (A and B) The solid box represents the known ceph-clav and paralogue gene clusters and the lines represent the unsequenced *S. clavuligerus* chromosome in the region flanking the respective gene clusters. The relative orientations and the identities of the cosmids are also indicated and the arrows indicate the ends of the cosmid inserts used as probes for chromosome walking. The filled circles indicate the sequenced T7 and T3 ends of cosmids, which have not been mapped. Therefore, their orientation with regard to the other cosmids shown is not known.

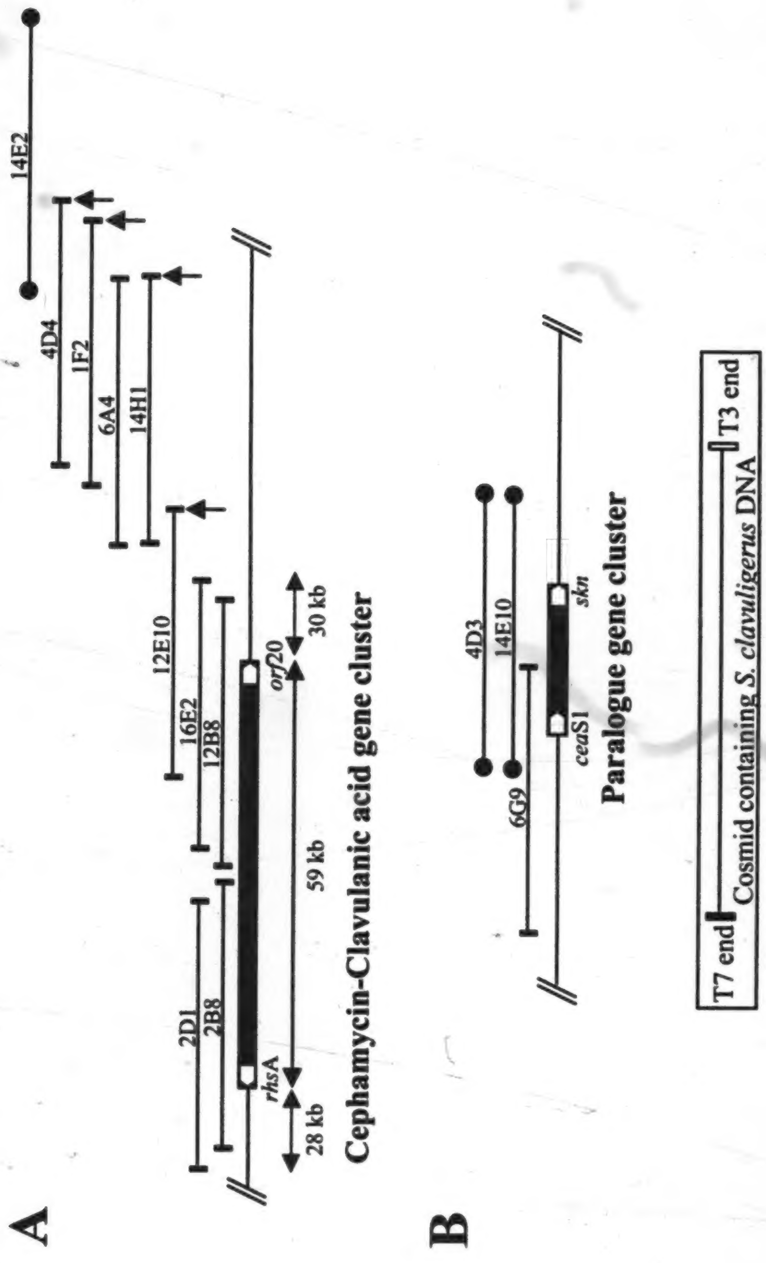

A
14E2
4D4
1F2
6A4
14H1
12E10
16E2
12B8
2D1
2B8
rhsA
orf20
28 kb
59 kb
30 kb
Cephamycin-Clavulanic acid gene cluster
B
4D3
14E10
6G9
ceaS1
skn
Paralogue gene cluster
T7 end
Cosmid containing S. clavuligerus DNA
T3 end

Table 3.11: DNA sequence of T7 and T3 ends of cloned inserts from cosmids used in this study.

Probe	Cosmid Identified	Insert end sequenced[a] (T3 or T7)	Most similar protein[b]	Percent identity and E-value[c]	Location on the *S. coelicolor* chromosome (Mb)[d]	Location on the *S. avermitilis* chromosome (Mb)[d]
*rhs*A	2B8	T3	Aminoadipyl-cysteinyl-valine synthetase (cephamycin arm of the ceph-clav gene cluster)	NA	NA	NA
		T7	Hypothetical protein	34%, 3e-4	8.48 (SCO7653)	5.71 (SAV4691)
*rhs*A	2D1	T3	Hypothetical protein	34%, 7e-05	5.02 (SCO4601)	3.17 (SAV2581)
		T7	Aminoadipyl-cysteinyl- valine synthetase (cephamycin arm of the ceph-clav gene cluster)	NA	NA	NA
*orf*20	12E10	T3	Putative integral membrane cell-cycle protein	68%, 1e-48	5.77 (SCO5302)	3.70 (SAV2951)
		T7	Orf13 (clavulanic arm of the ceph-clav gene cluster)	NA	NA	NA
*orf*20	12B8	T3	Isopenicillin N synthase (cephamycin arm of the ceph-clav gene cluster)	NA	NA	NA
		T7	Putative RNA polymerase ECF-subfamily sigma factor	74%, 7e-06	0.98 (SCO0942)	8.69 (SAV7292)

Probe	Cosmid Identified	Insert end sequenced[a] (T3 or T7)	Most similar protein[b]	Percent identity and E-value[c]	Location on the *S. coelicolor* chromosome (Mb)[d]	Location on the *S. avermitilis* chromosome (Mb)[d]
*orf*20	16E2	T3	Putative membrane protein	35%, 0.002	3.22 (SCO2963)	NIL
		T7	Carboxyethyl arginine synthase (clavulanic end of the ceph-clav gene cluster)	NA	NA	NA
12E10 T3 end	14H1	T3	Putative lon class III heat-shock ATP-dependent protease	98%, 8e-36	5.75 (SCO5285)	3.71 (SAV2966)
		T7	Conserved hypothetical protein	65%, 1e-30	5.78 (SCO5308)	3.69 (SAV2946)
12E10 T3 end	6A4	T3	Putative lon class III heat-shock ATP-dependent protease	96%, 2e-59	5.75 (SCO5285)	3.71 (SAV2966)
		T7	Conserved hypothetical protein	58%, 3e-08	5.78 (SCO5308)	3.69 (SAV2946)
14H1 T3 end	1F2	T3	Putative dehydrogenase	-	NIL	NIL
		T7	Hypothetical protein	-	NIL	NIL
14H1 T3 end	4D4	T3	Putative membrane protein	75%, 7e-25	2.19 (SCO2041)	7.42 (SAV6173)
		T7	Putative transcriptional regulator	81%, 1e-54	7.54 (SCO6784)	7.44 (SAV6198)

Probe	Cosmid Identified	Insert end sequenced[a] (T3 or T7)	Most similar protein[b]	Percent identity and E-value[c]	Location on the *S. coelicolor* chromosome (Mb)[d]	Location on the *S. avermitilis* chromosome (Mb)[d]
1F2 T3 and 4D4 T7 end	14E2	T3	Hypothetical protein	35%, 0.13	0.72 (SCO0683)	2.56 (SAV2100)
		T7	Integral membrane/ Gas vesicle synthesis proteins	27%, 0.078	2.77 (SCO2567)	0.74 (SAV593)
*pah*1	6G9	T3	Hypothetical protein	31%, 3e-07	1.48 (SCO1402)	4.07 (SAV4.07)
		T7	C6P, Putative aminotransferase (Paralogue gene cluster)	NA	NA	NA
*pah*1	14E10	T3	Putative membrane / Hypothetical proteins	38%, 0.18	2.17 (SCO2028)	7.79 (SAV6509)
		T7	Putative succinyltransferase/ Hypothetical protein	33%, 0.004	7.25 SCO0424	0.43 SAV6022
*cea*S1	4D3	T3	Conserved hypothetical protein	79%, 4e-05	0.64 (SCO0604)	0.9 (SAV778)
		T7	Putative transcriptional regulator	46%, 0.059	6.6 (SCO6020)	4.84 (SAV3919)

[a] Ends of the cloned inserts close to either the T3 or the T7 priming sites in the cosmid MCS

[b] Similarities at the amino acid level were determined by searching the *S. coelicolor* and *S. avermitilis* genome databases using the BLASTp online program

[c] Best E-values obtained are shown

[d] Respective location and locus tag of most similar protein on the *S. coelicolor* and *S. avermitilis* chromosomes

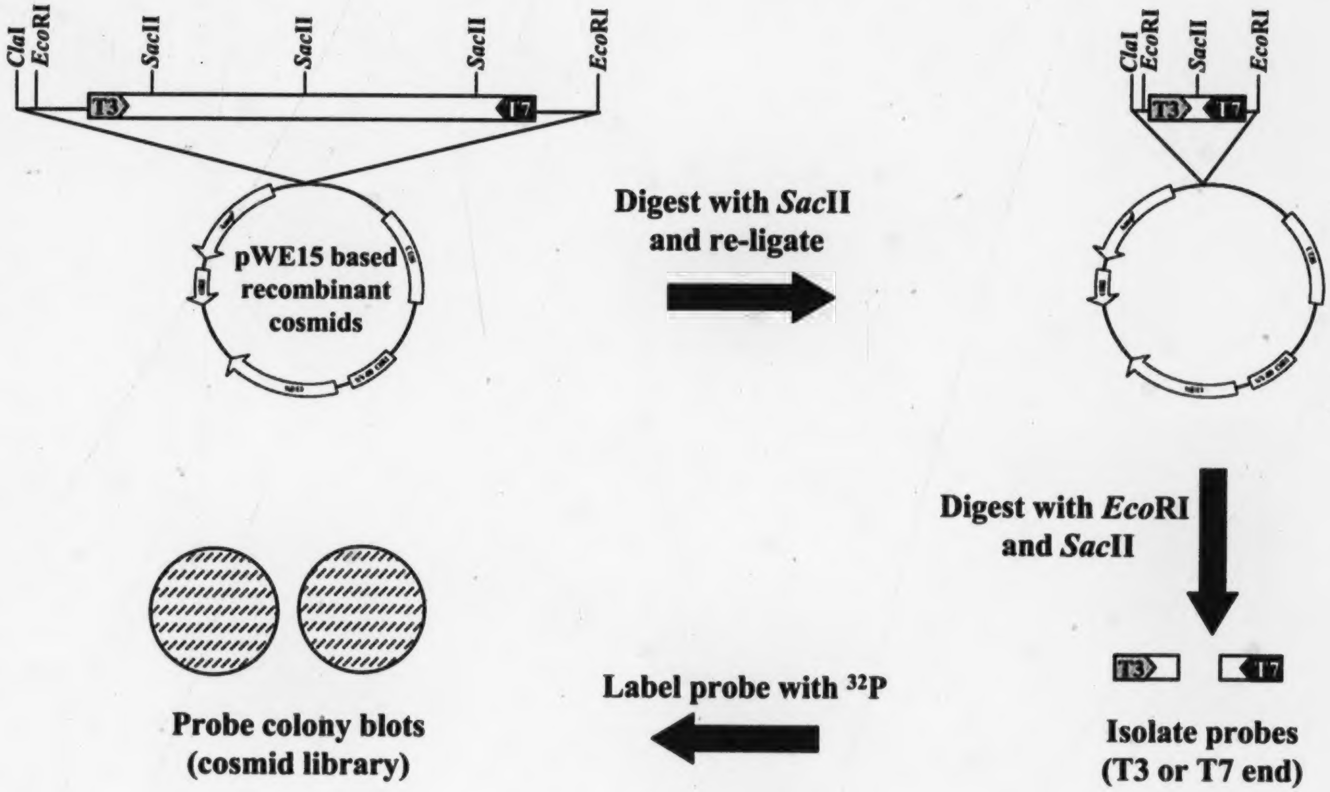
NA, not applicable as these cosmid ends carry genes that are known to encode proteins from the ceph-clav gene cluster

NIL, No significant similarity to any known or putative protein from either *S. coelicolor* or *S. avermitilis* was observed cloned insert from each cosmid, leaving only short fragments of *S. clavuligerus* genomic DNA from each end of the original insert (Figure 3.5.3). These short fragments were flanked by T3 and T7 promoter sequences within the cosmid vector, which were used to indicate the end of the insert from which they arose (Figure 3.5.3). Using this method, the plasmids p12E10-self, p14H1-self, p1F2-self and p4D4-self were prepared, and were used to isolate probes for the first, second, third and fourth consecutive walks, respectively. The short end fragments were released from the plasmids by *Sac*II/*Eco*RI digestion and were used as probes to re-screen the cosmid library in order to identify cosmids extending further downstream from the ceph-clav gene cluster (Figure 3.5.2). Cosmids isolated by this process were re-analyzed by Southern hybridization using the same probe to confirm the linkage and directionality between the isolated cosmids. No cosmids hybridizing to the paralogue gene cluster (Table 3.11) were found after four consecutive walks downstream from the clavulanic arm of the ceph-clav supercluster (Figure 3.5.2), indicating that the paralogue gene cluster is not located immediately downstream of the clavulanic acid gene cluster.

The region of the *S. clavuligerus* chromosome immediately upstream of the cephamycin end of the ceph-clav supercluster was also examined for the presence of the paralogue gene cluster (Figure 3.5.1). Cosmids including *rhs*A from the cephamycin end of the ceph-clav supercluster were isolated by screening the cosmid library with a probe consisting of an 800 bp *Eco*RI/*Bam*HI fragment from pDA115, which contains the 3′ end of *rhs*A and the 5′ end of *scl*U along with the intergenic region (Figure 3.5.1). Again, none of the cosmids containing the paralogue gene cluster (Table 3.11) were found to hybridize to cosmids carrying DNA from the region adjoining the cephamycin arm of the ceph-clav super cluster (Figure 3.5.2), indicating that the paralogue gene cluster is not located in this region.

The ends of all of the inserts in the cosmids isolated during chromosome walking studies were sequenced (Table 3.11) using the custom DNA oligonucleotide primers KTA2 and KTA3 (Table 2.4), which are based on sequences within the cosmid pWE15, near the T7 and the T3 junctions, respectively, where cloned DNA fragments would be inserted (Figure 3.5.3). Similarity searches were done to analyze the nature of the open reading frames located in these regions flanking the ceph-clav supercluster and the

Figure 3.5.3 Probe preparation for chromosome walking. The hollow box represents the regions of the *S. clavuligerus* chromosome present in the recombinant cosmids used in the analysis, and only the restriction sites used for preparing double stranded DNA probes are shown. The gray and black arrows represent the T3 and T7 DNA oligonucleotide primer-binding sequences respectively, used to determine the end of the cloned DNA insert in the cosmid, from which the probe was derived. The circular cosmid vector backbones are shown for reference only (Diagram is not to scale).


ClaI
EcoRI
SacII
SacII
SacII
EcoRI
T3
T7
pWE15 based recombinant cosmids
Digest with SacII and re-ligate
ClaI
EcoRI
SacII
EcoRI
T3
T7
Digest with EcoRI and SacII
T3
T7
Isolate probes (T3 or T7 end)
Label probe with ³²P
Probe colony blots (cosmid library)

paralogue gene cluster (Table 3.11). DNA sequence analysis did not indicate the presence of any of the known genes from the paralogue gene clusters in the region flanking the ceph-clav supercluster. Instead, a variety of ORFs were found, some showing high levels of similarity to putative proteins from both *S. coelicolor* and *S. avermitilis* (Table 3.11). In particular, several of the gene products encoded by genes from just beyond the clavulanic acid end of the ceph-clav supercluster, showed very high similarity to proteins encoded by genes at the 5.7-6.2 Mb region of the *S. coelicolor* chromosome (the corresponding region is 3.1-3.7 Mb in *S. avermitilis* due to reverse numbering of the genomic DNA sequence). However, other less highly conserved proteins from this region also showed similarity to *S. coelicolor* and *S. avermitilis* proteins spread more widely across the chromosome (Table 3.11). In addition, the predicted ORFs found in the region surrounding the paralogue gene cluster did not show any similarity to genes from the ceph-clav supercluster or the clavam gene cluster, and their most similar counterparts in *S. coelicolor* and *S. avermitilis* were spread widely across the chromosomes of the respective organisms (Table 3.11).

3.5.2 PFGE and Southern analysis

Due to the lack of any evidence from chromosome walking studies suggesting that the clavam or the paralogue gene clusters were located in the immediate vicinity of the ceph-clav supercluster, additional work was carried out to determine it there was a possibility of a more remote linkage. Studies were also performed to determine if the clavam and the paralogue gene clusters might be linked to each other, but not to the ceph-clav gene cluster. *S. clavuligerus* chromosomal DNA was digested with restriction enzymes, fractionated by PFGE and subjected to Southern analysis using probes specific for the respective gene clusters. The *orf*20 and *cvm*1 probes specific for the clavulanic acid and the clavam gene clusters respectively, were the same as those used for the isolation of cosmids containing DNA from the respective gene clusters (Section 3.5.1). A 796 bp *Sal*I fragment from p800-Sal, which includes the intergenic region between *pah*1 and *oat*1 as well as short stretches from the 3′ ends of both genes, was used as the probe for the paralogue gene cluster (Figure 3.5.1). Given the existence of the *pah*2 and *oat*2 genes within the ceph-clav gene cluster, care was taken in probe selection to ensure that cross-hybridization did not occur. The intergenic region between *pah*1 and *oat*1 was particularly suitable for this purpose because this region is not represented in the ceph-clav gene cluster (Figure 3.5.1).

Experiments were repeated at least twice using various restriction endonucleases that cleave the *Streptomyces* chromosome at low frequencies. The entire ceph-clav, clavam and paralogue gene clusters were examined by DNA sequence analysis to confirm that there were no cleavage sites present within the respective gene clusters for any of the restriction endonucleases used. The only exception was the *pcb*AB gene from the ceph-clav super cluster, which has been partially sequenced only. To ensure *pcb*AB did not contain any sites for the enzymes used in this study, the *pcb*AB containing recombinant cosmid 8-51 (Table 2.3) was subjected to restriction analysis using the respective enzymes, none of which cut within the insert in 8-51 (Figure 3.5.4).

Using this procedure, high molecular weight restriction fragments encoding each of the gene clusters were detected, but no two probes hybridized to a single DNA fragment (Figure 3.5.5, Table 3.12). The very different sizes of the restriction fragments that hybridized to the probes specific for the three gene clusters indicated that the three clavam metabolite-associated gene clusters must be separated from one another by considerable distances (Table 3.12).

Although there are no restriction sites present within any of the three respective gene clusters for the enzymes used in this study, there is an *Hin*dIII site located downstream of the *cvm*6 gene in the clavam gene cluster (Jensen *et al.* Unpublished) (Figure 3.5.1). To ensure that this *Hin*dIII restriction site was not the delineating boundary between any of the gene clusters being subjected to analysis, the 3.6 kb *Kpn*I/*Sac*I DNA fragment from pJENII, which spans the *Hin*dIII site in question (Figure 3.5.1), was used as the probe to screen *Hin*dIII digested *S. clavuligerus* chromosomal DNA by Southern hybridization. The 3.6 kb probe hybridized to two *Hin*dIII DNA fragments, which corresponded to the 190 kb fragment from the clavam gene cluster (Table 3.12), and to a smaller (<100 kb) fragment, but not to any other DNA fragment from the ceph-clav or paralogue gene clusters (Figure 3.5.5). This indicated that this *Hin*dIII site in question was not responsible for separating two gene clusters that may have been originally linked.

Figure 3.5.4 Restriction analysis of *pcb*AB containing recombinant cosmid 8-51. The pLAFR3 based recombinant cosmid 8-51 was digested with the restriction enzymes used in PFGE and linkage analysis, to determine if the cloned insert 8-51 contained any sites for the respective enzymes. Lanes corresponding to the cosmid vector without any insert and the recombinant cosmid 8-51 are shown side by side. The restriction enzymes used in the analysis are also indicated and lane M was λ phage DNA digested with *Bst*EII, and was used as the marker for size estimation.

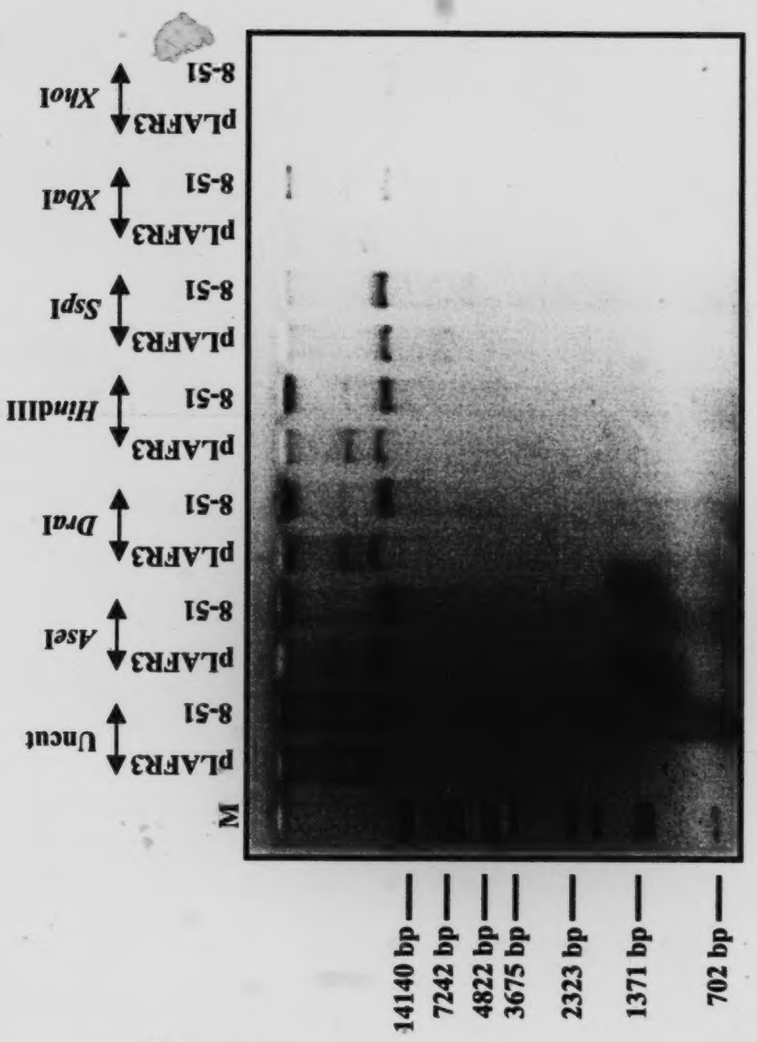

M
pLAFR3
8-S1
Uncut
pLAFR3
8-S1
AseI
pLAFR3
8-S1
DraI
pLAFR3
8-S1
HindIII
pLAFR3
8-S1
SspI
pLAFR3
8-S1
XbaI
pLAFR3
8-S1
XhoI
14140 bp
7242 bp
4822 bp
3675 bp
2323 bp
1371 bp
702 bp

Figure 3.5.5 PFGE and Southern analysis of *S. clavuligerus* chromosomal DNA. (A) Yeast chromosome PFG markers were fractionated by PFGE and used for size estimation (B) *S. clavuligerus* chromosomal DNA, undigested or digested with *Hin*dIII was fractionated by PFGE. The sizes of the bands corresponding to the two GLPs (460 and 120 kb) from *S. clavuligerus* are shown in the lane containing undigested DNA. (C) Chromosomal DNA, undigested and digested with *Hin*dIII was analyzed by Southern hybridization using probes specific for the clavulanic acid arm of the ceph-clav gene cluster, the clavam and the paralogue gene clusters respectively (described in results). The sizes of the DNA fragments (in kb) hybridizing to the respective probes are shown. (D) Chromosomal DNA digested with *Hin*dIII was analyzed by Southern hybridization using the 3.6 kb *KpnI/Sac*I fragment from pJENII as the probe, which spanned the *Hin*dIII restriction site present downstream of the clavam gene cluster (see Figure 3.5.1).

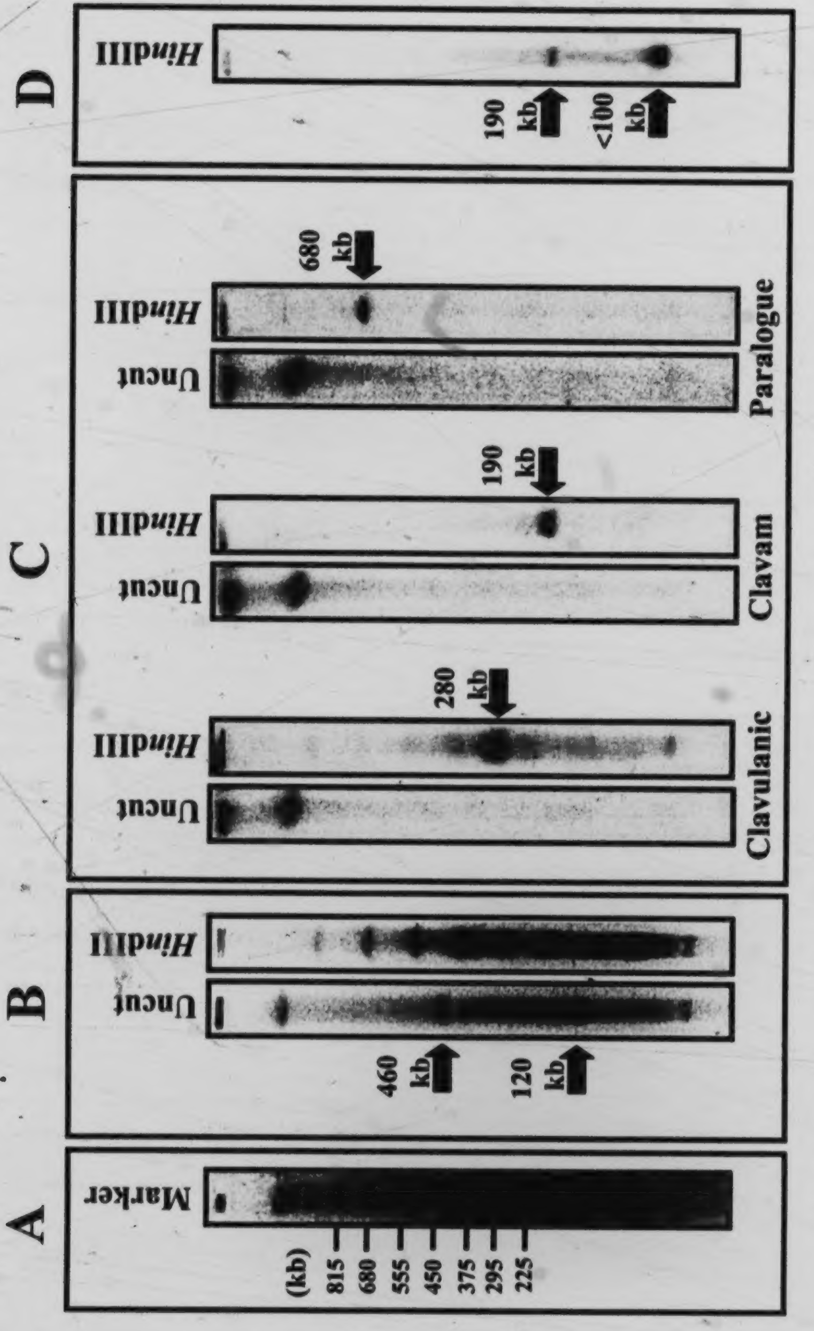

A
Marker
(kb)
815
680
555
450
375
295
225
B
Uncut
HindIII
460 kb
120 kb
C
Uncut
HindIII
280 kb
Clavulanic
Uncut
HindIII
190 kb
Clavam
Uncut
HindIII
680 kb
Paralogue
D
HindIII
190 kb
<100 kb

Table 3.12: Approximate sizes of restriction fragments hybridizing to probes specific for the clavulanic acid, the clavam and the paralogue gene clusters.

Restriction enzyme[a]	Size of the restriction fragment that hybridized to the probe specific for the respective gene clusters (kb)[b]		
	Clavulanic	Clavam	Paralogue
*Ase*I	Nil[c]	350	915
*Bfr*I	375	50	815
*Dra*I	560	630	295
*Hin*dIII	290	190	680
*Ssp*I	760	225	480
*Xba*I	450	870	610
*Xho*I	310	Nil[c]	910

[a] Restriction endonucleases used for *en bloc* digestion of wild type *S. clavuligerus* chromosomal DNA before PFGE

[b] Probes specific for each gene cluster are described in Materials and Methods

[c] Weak hybridization signals were observed from high molecular weight restriction fragments, which were not resolved under the conditions used

S. clavuligerus also possesses three linear plasmids that are 460 kb, 120 kb and 11.7 kb in size. The two large plasmids are often referred to as Giant Linear Plasmids (GLPs) in *S. clavuligerus*. The 11.7 kb pSCL1 (Wu and Roy, 1993) and the 120 kb pSCL2 (Wu, 2003) plasmids have been sequenced (~80% of pSCL2 has been sequenced) and do not encode any proteins involved in β-lactam metabolite biosynthesis. It has also been shown by Southern hybridization that the two GLPs do not encode the genes involved in cephamycin C biosynthesis (Netolitzky *et al.*, 1995). The results presented here further indicate that none of the three gene clusters involved in clavam metabolite biosynthesis are encoded by the two GLPs, as these plasmids did not hybridize to any of the three gene cluster-specific probes (Figure 3.5.5).

Overall, results from linkage studies indicate that the ceph-clav, the clavam and the paralogue gene clusters do not seem to be physically linked to form a still larger gene cluster together, as hypothesized earlier.

3.6 Regulation of *cea*S1 and *cea*S2 expression

When *S. clavuligerus* mutants defective in *cea*S1 and *cea*S2 were prepared individually, both mutant strains still retained some ability to produce both clavulanic acid and the 5*S* clavams, depending on the fermentation medium used (Sections 3.1 and 3.3). The *cea*S1 mutant produced both clavulanic acid and the 5*S* clavams in soy medium, but only clavulanic acid in SA medium. In contrast, the *cea*S2 mutant occasionally (and unpredictably) produced small amounts of clavulanic acid and the 5*S* clavam metabolites in soy medium only, while no clavulanic acid or 5*S* clavam production was observed in cultures grown in SA medium. Based on the observed phenotypes, it was postulated that *cea*S1 was expressed in soy medium only, whereas *cea*S2 was expressed in both soy and SA media. To verify this hypothesis, the effect of growth in soy and SA media on both *cea*S1 and *cea*S2 expression was examined at the transcriptional level. Although it was known that the cephamycin-clavulanic acid pathway-specific regulator CcaR regulates the expression of the late clavulanic acid biosynthetic genes through ClaR (Paradkar *et al.*, 1998), which is the pathway-specific transcriptional regulator for clavulanic acid only, the effect of CcaR on *cea*S1 and *cea*S2 expression was not known. Therefore, the effect of CcaR on the transcription of *cea*S1 and *cea*S2 was also examined in this study. Note: portions of this section have been published as part of the manuscript by Tahlan *et al.* (2004a).

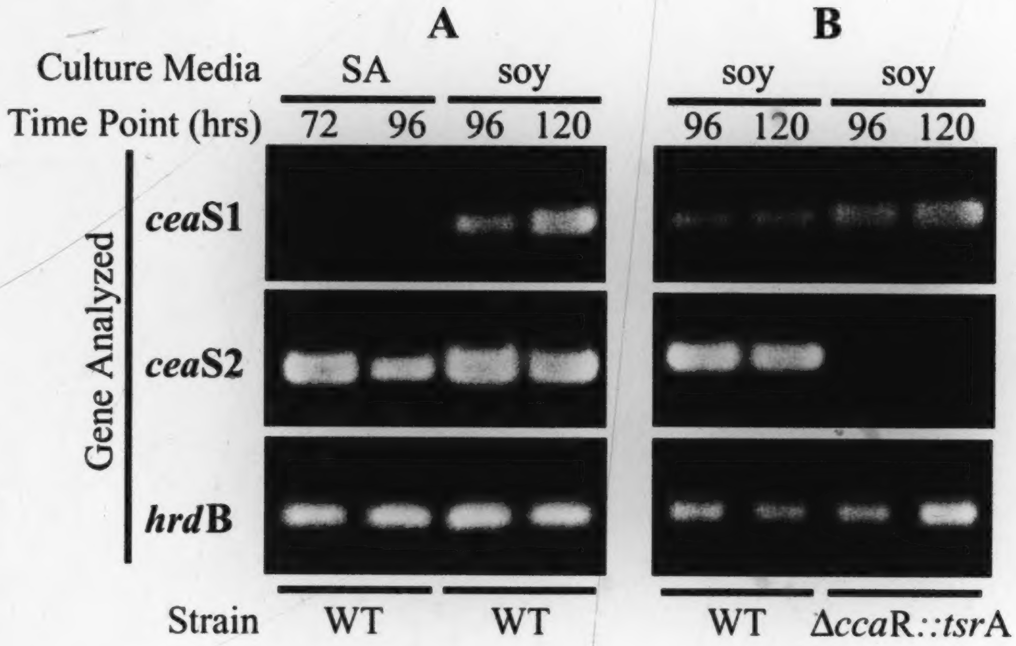
3.6.1 Nutritional regulation of *cea*S1 and *cea*S2 expression

The nutritional regulation of *cea*S1 and *cea*S2 expression was analyzed by reverse transcriptase-PCR (RT-PCR) and by the use of the enhanced green fluorescent protein-encoding gene (e*gfp*) as a reporter, as described in the following Sections.

3.6.1.1 RT-PCR analysis

DNA oligonucleotide primers used in RT-PCR analysis were carefully chosen to prevent cross-amplification of the *cea*S genes. The reverse primers ceaS1-RT-Rev, ceaS2-RT-Rev and CAN 122 (Table 2.4) were used to synthesize cDNA corresponding to the *cea*S1, *cea*S2 and *hrd*B transcripts, respectively. The *hrd*B gene encodes a constitutively expressed sigma factor in *Streptomyces* (Buttner *et al.*, 1990; Kormanec and Farkasovsky, 1993; Kormanec *et al.*, 1992; Marcos *et al.*, 1995; Shinkawa *et al.*, 1995), and its transcription was monitored as a control. The primer pairs ceaS1-RT-For/ceaS1-RT-Rev, ceaS2-RT-For/ceaS2-RT-Rev and CAN123/CAN122 were used for the PCR amplification of the 433 bp *cea*S1, the 452 bp *cea*S2 and the 368 bp *hrd*B RT-products, respectively. The *cea*S1 and the *cea*S2 RT-PCR products spanned the regions corresponding to +292 to + 725 bp and +110 to +562 bp relative to the putative start codons of *cea*S1 and *cea*S2, respectively. RNA isolated from wild type *S. clavuligerus* grown on SA medium for 72 and 96 hours, and soy medium for 96 and 120 hours was used in the analysis. When RNA isolated from SA cultures was subjected to analysis, *cea*S1 transcripts were not detected, whereas the same samples showed the presence of *cea*S2 transcripts (Figure 3.6.1). On similar analyses of RNA isolated from soy grown cultures, both *cea*S1 and *cea*S2 transcripts were detected by RT-PCR (Figure 3.6.1). In addition, *hrd*B transcripts were detected at similar levels in all samples tested (Figure 3.6.1). HPLC analysis of culture supernatants showed expected levels of clavulanic acid and 5*S* clavams in cultures grown in both media used for RNA isolation and analysis (data not shown).

Figure 3.6.1 Assessment of *cea*S1 and *cea*S2 transcript levels by RT-PCR. RNA samples isolated from various strains of *S. clavuligerus* were analyzed by RT-PCR using primers specific for *cea*S1, *cea*S2 and *hrd*B. (A) Analysis of RNA from *S. clavuligerus* wild type (WT) cultures grown on SA medium for 72 and 96 hours, and from cultures grown on soy medium for 96 and 120 hours. (B) Analysis of RNA from *S. clavuligerus* wild type cultures grown on soy medium for 96 and 120 hours, and *S. clavuligerus* Δ*cca*R::*tsr*A mutant cultures grown on soy medium for 96 and 120 hours.


A
B
Culture Media
SA
soy
soy
soy
Time Point (hrs)
72
96
96
120
96
120
96
120
Gene Analyzed
*cea*S1
*cea*S2
*hrd*B
Strain
WT
WT
WT
Δ*cca*R::*tsr*A

3.6.1.2 Use of *egfp* for promoter activity analysis

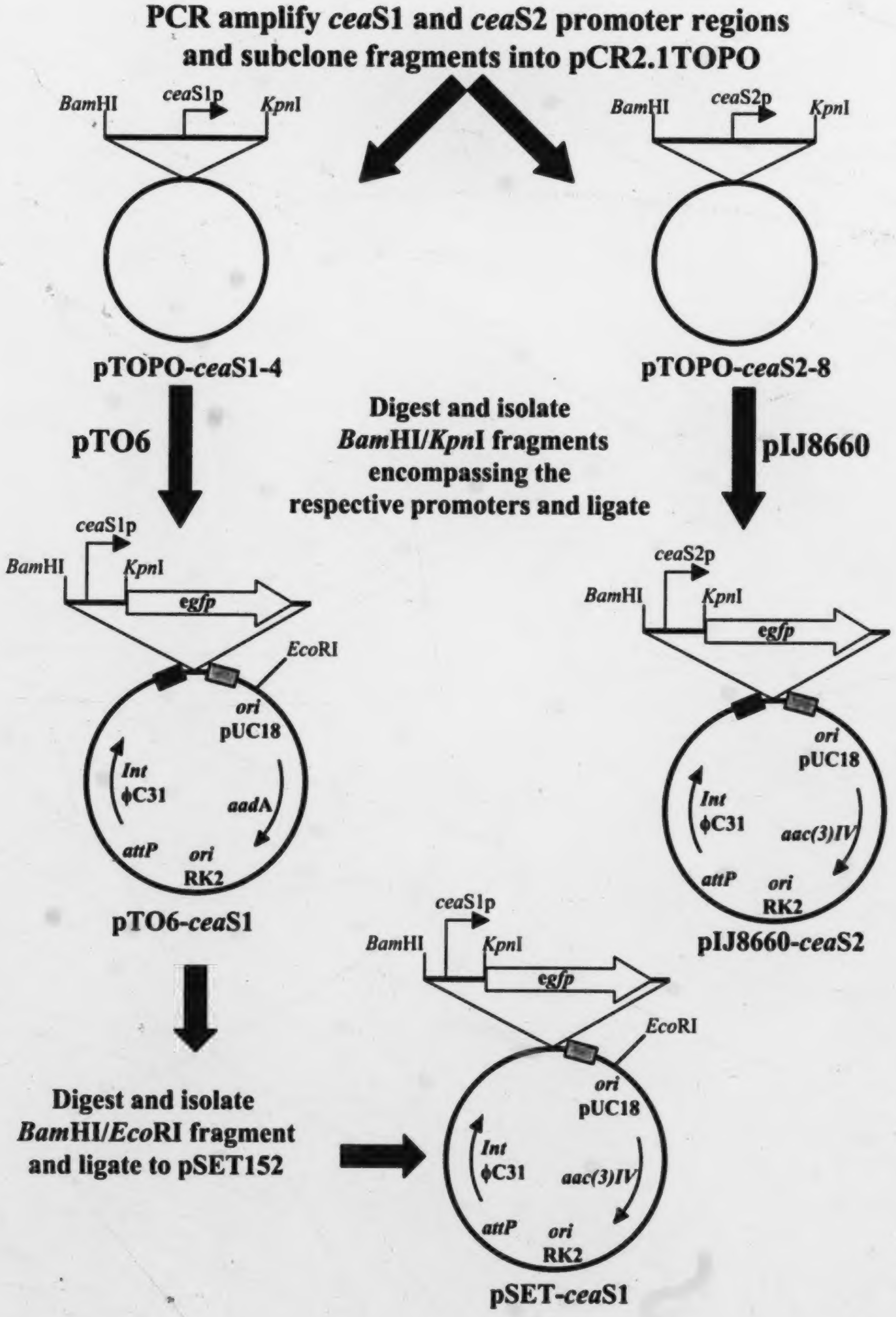
To confirm the results obtained by RT-PCR, the promoter regions of *cea*S1 and *cea*S2 were also analyzed in an *egfp* promoter probe system. A 781 bp DNA fragment spanning the *cea*S1 promoter region was amplified by PCR using p2.8-18 as the template and the primer pair KTA-ceaS1-For and KTA-ceaS1-Rev. Similarly, a 721 bp DNA fragment encompassing the *cea*S2 promoter region was amplified by PCR using pBB5.3A as the template and the primers KTA-ceaS2-For and KTA-ceaS2-Rev. The PCR products were treated with Taq DNA polymerase before ligation to pCR2.1TOPO as per the manufacturer's instructions (Invitrogen). This gave pTOPO-*cea*S1-4 and pTOPO-*cea*S2-8, which contained the *cea*S1 and the *cea*S2 promoter regions in pCR2.1TOPO, respectively. Double stranded DNA sequence of the inserts was obtained using universal primers to ensure that no mutations were introduced during PCR amplification.

The *cea*S2 promoter region from pTOPO-*cea*S2-8 was isolated as a *Bam*HI/*Kpn*I fragment and ligated into the corresponding sites of pIJ8660 to give pIJ8660-ceaS2. pIJ8660-ceaS2 contained the *cea*S2 promoter region fused to a promoterless *egfp* gene for use as a reporter of expression driven by the *cea*S2 promoter (Figure 3.6.2).

For unexplained reasons, the *cea*S1 promoter region could not be subcloned into pIJ8660, and therefore an alternative approach was used. The *cea*S1 promoter region from pTOPO-*cea*S1-4 was isolated as a *Bam*HI/*Kpn*I fragment and ligated into the corresponding sites of pTO6 to give pTO6-*cea*S1. The entire cassette encompassing the *cea*S1 promoter region fused to the *egfp* gene from pTO6-ceaS1 was isolated as a 2.37 kb *Bam*HI/*Eco*RI fragment and introduced into the corresponding sites of pSET152 to give pSET-*cea*S1, which served as the *cea*S1 reporter construct (Figure 3.6.2).

The plasmids pSET-*cea*S1 and pIJ8660-*cea*S2 containing the promoter regions of *cea*S1 and *cea*S2 (-501 to + 215 and -551 to + 95 bp relative to the putative start codons of *cea*S1 and *cea*S2, respectively) in front of a promoter-less *egfp* gene, were introduced into wild type *S. clavuligerus* by conjugation. The plasmid pIJ8660 containing the promoterless *egfp* gene was also introduced into the wild type strain (Table 2.2) and functioned as the control for determination of and normalizing for background levels of EGFP expression. Strains that had the respective plasmids integrated at the ϕC31 *attB* site

Figure 3.6.2 Preparation of the *cea*S1 and *cea*S2 promoter-e*gfp* reporter constructs. The reporter plasmids pSET-*cea*S1 and pIJ8660-*cea*S2, which contained the *cea*S1 and *cea*S2 promoters fused to a promoterless enhanced green fluorescence protein (e*gfp*) gene respectively, were used to examine transcription driven by the respective promoters, by measuring fluorescence due to EGFP expression/excitation. A schematic showing the strategy used to prepare plasmids pSET-*cea*S1 and pIJ8660-*cea*S2 is outlined. The large circles represent plasmid vector backbones and the hollow arrows represent the e*gfp* gene with the arrowhead indicating its orientation. The bent arrows represent the *cea*S1 and *cea*S2 promoters (*cea*S1p and *cea*S2p, respectively). The plasmid origins of replication (*ori*), gene encoding the phage ϕC31 integrase (*int* ϕC31), site in the vector for integrating into the chromosome (*attP*), the apramycin resistance gene [*aac(3)IV*] and restriction sites used in the preparation of the respective plasmids are also shown. The black and gray boxes represent the *tfd* (major transcriptional terminator of phage fd) and the *to* (transcriptional terminator from phage λ), respectively.


PCR amplify *cea*S1 and *cea*S2 promoter regions and subclone fragments into pCR2.1TOPO
*Bam*HI
*cea*S1p
*Kpn*I
*Bam*HI
*cea*S2p
*Kpn*I
pTOPO-*cea*S1-4
pTOPO-*cea*S2-8
pTO6
Digest and isolate *Bam*HI/*Kpn*I fragments encompassing the respective promoters and ligate
pIJ8660
*cea*S1p
*Bam*HI
*Kpn*I
egfp
*Eco*RI
ori pUC18
Int ϕC31
aadA
attP
ori RK2
pTO6-*cea*S1
*cea*S2p
*Bam*HI
*Kpn*I
egfp
ori pUC18
Int ϕC31
aac(3)IV
attP
ori RK2
pIJ8660-*cea*S2
Digest and isolate *Bam*HI/*Eco*RI fragment and ligate to pSET152
*cea*S1p
*Bam*HI
*Kpn*I
egfp
*Eco*RI
ori pUC18
Int ϕC31
aac(3)IV
attP
ori RK2
pSET-*cea*S1

in the chromosome were isolated based on apramycin resistance (apr^r) and designated C1G and C2G. Fluorescence arising due to EGFP expression was used as a reporter to monitor transcription driven from the respective promoters. Due to high levels of autofluorescence and photobleaching, all measurements were non-quantative. The *S. clavuligerus* reporter strains C1G and C2G (Table 2.2) were grown on soy and SA media for 72 hours, after which mycelia were harvested and analyzed by confocal microscopy to examine fluoresence arising due to EGFP expression/excitation. Fluoresence was observed in all samples except in the C1G strain grown on SA medium (Figure 3.6.3). Cell free extracts (CFEs) were also prepared from the same samples that were subjected to confocal microscopy and the presence of EGFP in the CFEs was analyzed by western blotting. A 27 kDa band corresponding to EGFP was observed in all of the samples except for the lane containing CFE from the C1G strain grown on SA medium (Figure 3.6.3), confirming that the fluorescence observed in the samples was due to true EGFP expression.

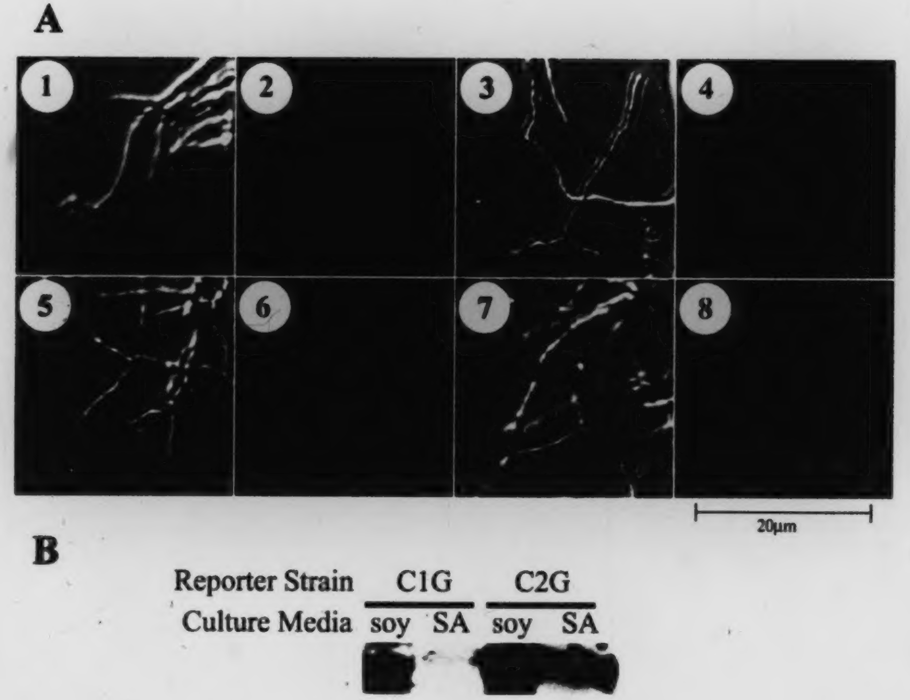
Overall, results from both RT-PCR and EGFP reporter analysis clearly indicated that *cea*S1 was transcribed in soy medium only, whereas *cea*S2 was transcribed in both soy and SA media.

3.6.2 Effect of CcaR on *cea*S1 and *cea*S2 transcription

RNA isolated from wild type and the Δ*cca*R::*tsr*A strains of *S. clavuligerus* grown on soy medium for 96 and 120 hours was analyzed by RT-PCR as described in Section 3.6.1, to monitor the effect of CcaR status on *cea*S1 and *cea*S2 expression. The transcription of *cea*S1 was comparable in both the wild type and the Δ*cca*R::*tsr*A strains (Figure 3.6.1). When *cea*S2 transcription was examined in the same RNA samples, almost no transcripts were detectable in the Δ*cca*R::*tsr*A mutant as compared to the wild type strain (Figure 3.6.1). The expression of *hrd*B was also monitored as a control and was found to be constant in all of the samples (Figure 3.6.1).

The levels of clavulanic acid and 5*S* clavam production in the wild type and the Δ*cca*R::*tsr*A mutant cultures used to isolate RNA were also determined by HPLC analysis of culture supernatants. The wild type strain produced expected levels of clavulanic acid and the 5*S* clavams after 96 and 120 hours of growth on soy medium.

Figure 3.6.3 Use of EGFP as a reporter to detect *cea*S1 and *cea*S2 promoter activity. (A) Mycelia from *S. clavuligerus* EGFP reporter strains grown on SA and soy media for 72 hours were analyzed by confocal microscopy. Both differential interference contrast (DIC, odd numbered panels) and fluorescence images (even numbered panels) were obtained and are shown side by side. Panels 1 and 2, *cea*S1 reporter grown on SA medium; panels 3 and 4, *cea*S2 reporter grown on SA medium; panels 5 and 6, *cea*S1 reporter grown on soy medium; panels 7 and 8, *cea*S2 reporter grown on soy medium. (B) Detection of EGFP in cell free extracts (CFEs) from *S. clavuligerus* EGFP reporter strains grown on soy and SA media for 72 hours. Cell free extracts from samples subjected to microscopic analysis were also analyzed by western blotting using commercially available antibodies raised against EGFP. C1G, lanes containing CFEs from the *cea*S1 reporter strain; C2G, lanes containing CFEs from the *cea*S2 reporter strain. The different culture media used to grow the reporter strains are also indicated.


A
1
2
3
4
5
6
7
8
20µm
B
Reporter Strain
C1G
C2G
Culture Media
soy
SA
soy
SA

Under the same conditions the Δ*cca*R*::tsr*A strain strain did not produce any detectable levels of clavulanic acid, whereas normal 5*S* clavam production was observed (Figure 3.6.4). This indicated that the transcription of *cea*S1 is not regulated by CcaR whereas that of *cea*S2 is, and that CcaR regulates the biosynthesis of clavulanic acid but not that of the 5*S* clavams.

3.7 Transcriptional mapping of the *cea*S1 and *cea*S2 promoters

Transcriptional analysis of *cea*S1 and *cea*S2 demonstrated that the two genes are differently regulated in terms of nutrition and the effect of CcaR on their transcription (Section 3.6). The transcriptional start points (TSPs) of both *cea*S1 and *cea*S2 were mapped by S1 nuclease protection and primer extension analyses to facilitate in the identification of promoters and potential regulatory elements, which may control the transcription of the respective genes. RNA used for the analysis was isolated from wild type *S. clavuligerus* grown on soy medium (96 and 120 hours). The primers ceaS1-S1-For and ceaS1-PR-EX along with p2.8-18 as template were used to prepare a 297 bp probe extending from -261 to +26 bp relative to the putative *cea*S1 start codon. S1 nuclease protection analysis using the 297 bp probe indicated that the *cea*S1 TSP was located 98-99 bp upstream of the *cea*S1 start codon (Figure 3.7.1A). Primer extension analysis was also conducted using the reverse primer ceaS1-PR-EX and RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 hours, confirming that the *cea*S1 TSP mapped to a C residue located 98 bp upstream of the *cea*S1 ATG start codon (Figures 3.7.1B).

The primers ceaS2-S1-For and ceaS2-PR-EX along with the template plasmid pBB5.3A were used to prepare a 236 bp probe extending from -204 to + 22 bp relative to the *cea*S2 start codon, which was used in an S1 nuclease protection assay to map the *cea*S2 transcriptional start point. The results indicated that the *cea*S2 TSP was located 51-52 bp upstream of the *cea*S2 start codon (Figure 3.7.2A). Primer extension analysis using the reverse primer ceaS2-PR-EX and RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 hours (Figure 3.7.2B), confirmed that the *cea*S2 transcript originated from a G residue located 51 bp upstream of the *cea*S2 ATG start codon (Figures 3.7.2 and 3.7.3).

Figure 3.6.4 Clavulanic acid and 5*S* clavams produced by *S. clavuligerus* wild type and Δ*cca*R::*tsr*A mutant strains. Supernatants from wild type and Δ*cca*R::*tsr*A mutant cultures grown in soy media for 96 (A) and 120 (B) hours were analyzed by HPLC. The peaks are as follows: Peak 1, clavaminic acid; peak 2, clavam-2-carboxylic acid; peak 3, 2-hydroxymethyl clavam; peak 4, clavulanic acid.

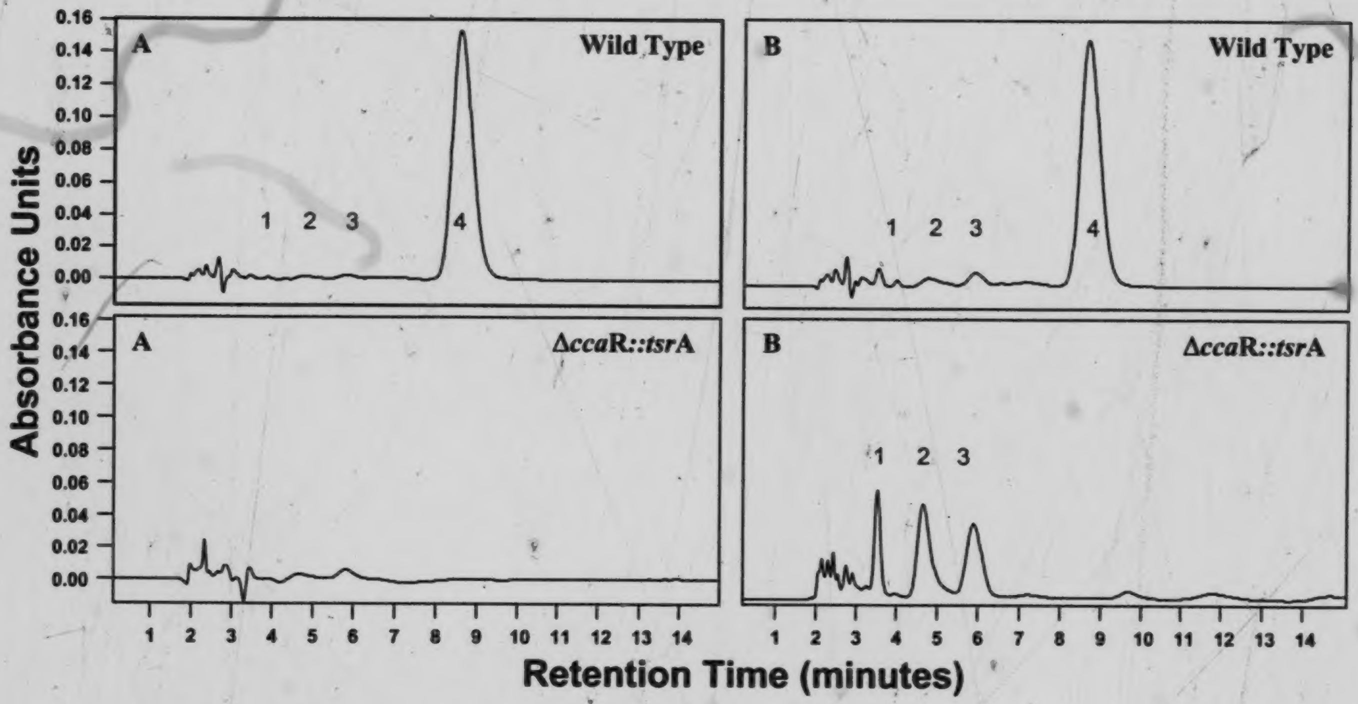

A
Wild Type
B
Wild Type
A
ΔccaR::tsrA
B
ΔccaR::tsrA
1
2
3
4
Absorbance Units
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Retention Time (minutes)

Figure 3.7.1 Mapping of the *cea*S1 transcriptional start point. (A) A 297 bp probe (-261 to +26 bp relative to the putative *cea*S1 start codon, plus 10 bp of non-homologous sequence) was used in the S1 protection assay. (B) Primer extension analysis using the *cea*S1 specific reverse primer, ceaS1-PR-EX. (A and B) The sequencing ladders were prepared using the reverse primer ceaS1-PR-EX along with p2.8-18 as template. Lanes G, A, T and C represent the sequencing ladder. (A) Lanes 1 and 2, RNA from wild type *S. clavuligerus* grown on soy medium for 96 and 120 hours and subjected to analysis; lane P+S1, control lane with unprotected probe digested with S1 nuclease; lane P, undigested probe control. (B) Lane 1, RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 hours and subjected to primer extension analysis. The most probable transcriptional start points are marked by (*).

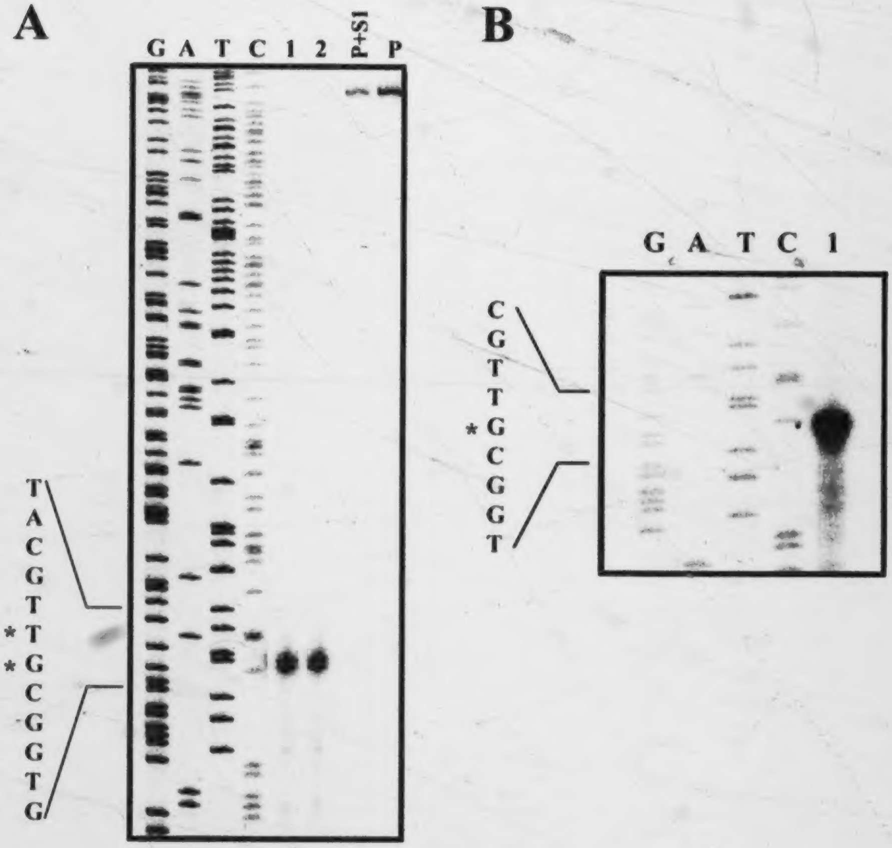

A
G A T C 1 2 P+S1 P
T
A
C
G
T
* T
* G
C
G
G
T
G
B
G A T C 1
C
G
T
T
* G
C
G
G
T

Figure 3.7.2 Mapping of the *cea*S2 transcriptional start point. (A) S1 nuclease protection analysis using a 236 bp probe (-204 to + 22 bp relative to the *cea*S2 start codon, plus 10 bp of non-homologous sequence), used to map the *cea*S2 transcriptional start point. (B) Primer extension analysis using the *cea*S2 specific reverse primer, ceaS2-PR-EX. (A and B) The DNA sequencing ladders were prepared for size estimation using the reverse primer ceaS2-PR-EX and template plasmid pBB5.3A. Lanes G, A, T and C represents the sequencing ladder (A) lanes 1 and 2, RNA from wild type *S. clavuligerus* grown on soy medium for 96 and 120 hours and subjected to analysis; lane P+S1, control lane with unprotected probe digested with S1 nuclease; lane P, undigested probe control. (B) Lane 1, RNA from wild type *S. clavuligerus* subjected to primer extension analysis. (A and B) The most probable transcriptional start points are marked by (*).

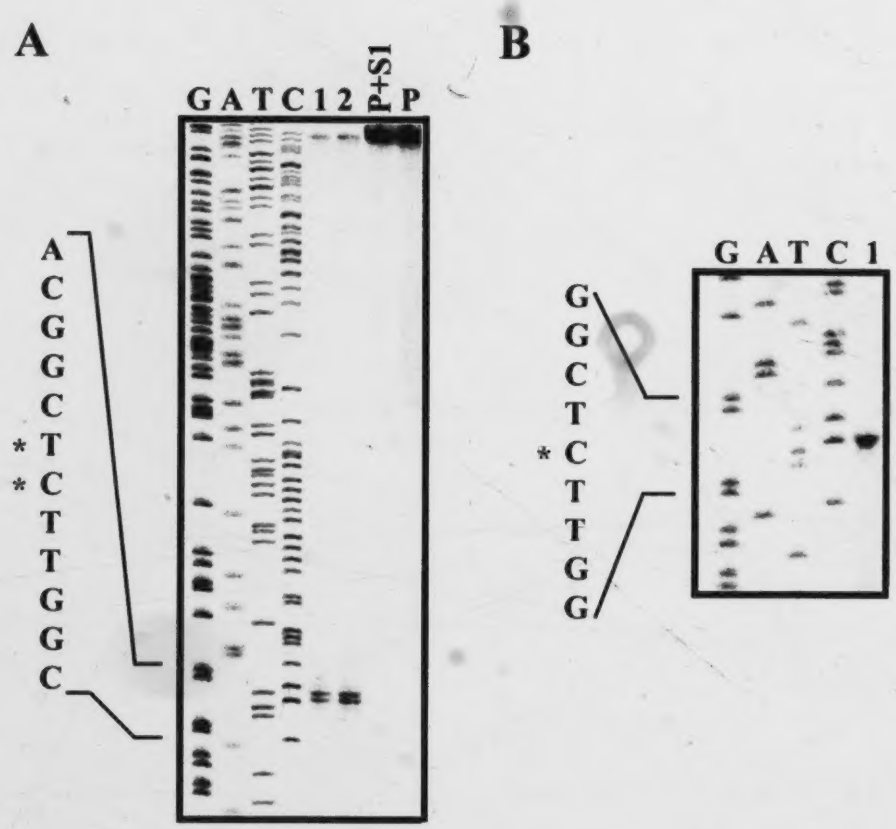
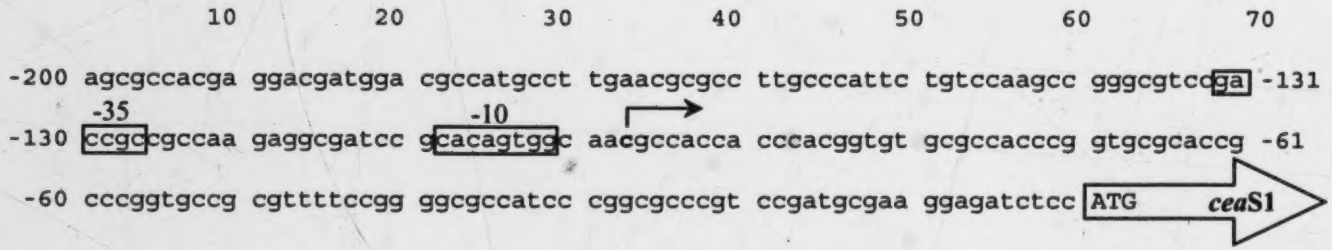
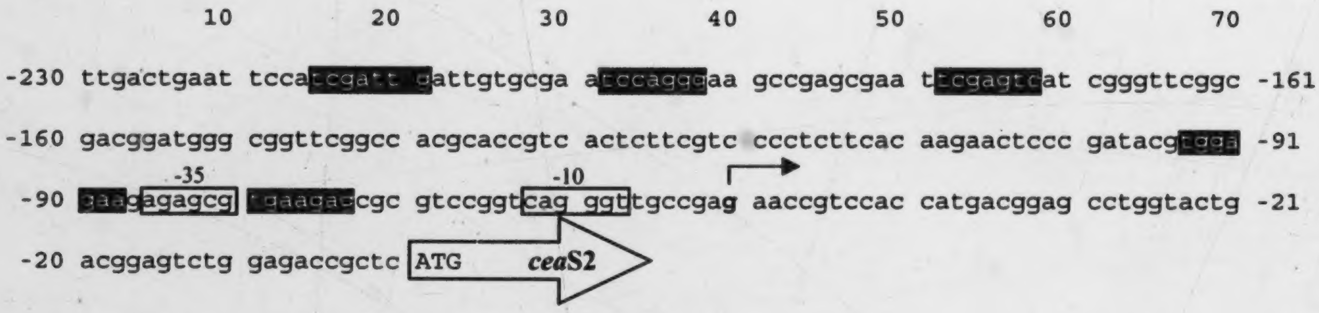

A
G A T C 1 2 P+S1 P
A
C
G
G
C
* T
* C
T
T
G
G
C
B
G A T C 1
G
G
C
T
* C
T
T
G
G

Figure 3.7.3 DNA sequence of the *cea*S1 and *cea*S2 promoter regions. (A and B) The hollow arrows represent the respective genes with the arrowhead representing the orientation of the gene. The fine arrow indicates the transcriptional start points and the respective –10 and –35 promoter regions are also shown. (A) DNA sequence of the *cea*S1 promoter region. (B) DNA sequence of the *cea*S2 promoter region showing the possible heptameric repeats, shown in filled boxes, which are recognized by *Streptomyces* antibiotic regulatory proteins.

A


10 20 30 40 50 60 70
-200 agcgccacga ggacgatgga cgccatgcct tgaacgcgcc ttgcccattc tgtccaagcc gggcgtccga -131
-35
-10
-130 ccgccgccaa gaggcgatcc gcacagtggc aacgccacca cccacggtgt gcgccacccg gtgcgcaccg -61
-60 cccggtgccg cgttttccgg ggcgccatcc cggcgcccgt ccgatgcgaa ggagatctcc
ATG
ceaS1

B


10 20 30 40 50 60 70
-230 ttgactgaat tccatcgatt gattgtgcga atccagggaa gccgagcgaa ttcgagtcat cgggttcggc -161
-160 gacggatggg cggttcggcc acgcaccgtc actcttcgtc ccctcttcac aagaactccc gatacgtgga -91
-35
-10
-90 gaagagagcg tgaagaccgc gtccggtcag ggttgccgag aaccgtccac catgacggag cctggtactg -21
-20 acggagtctg gagaccgctc
ATG
ceaS2

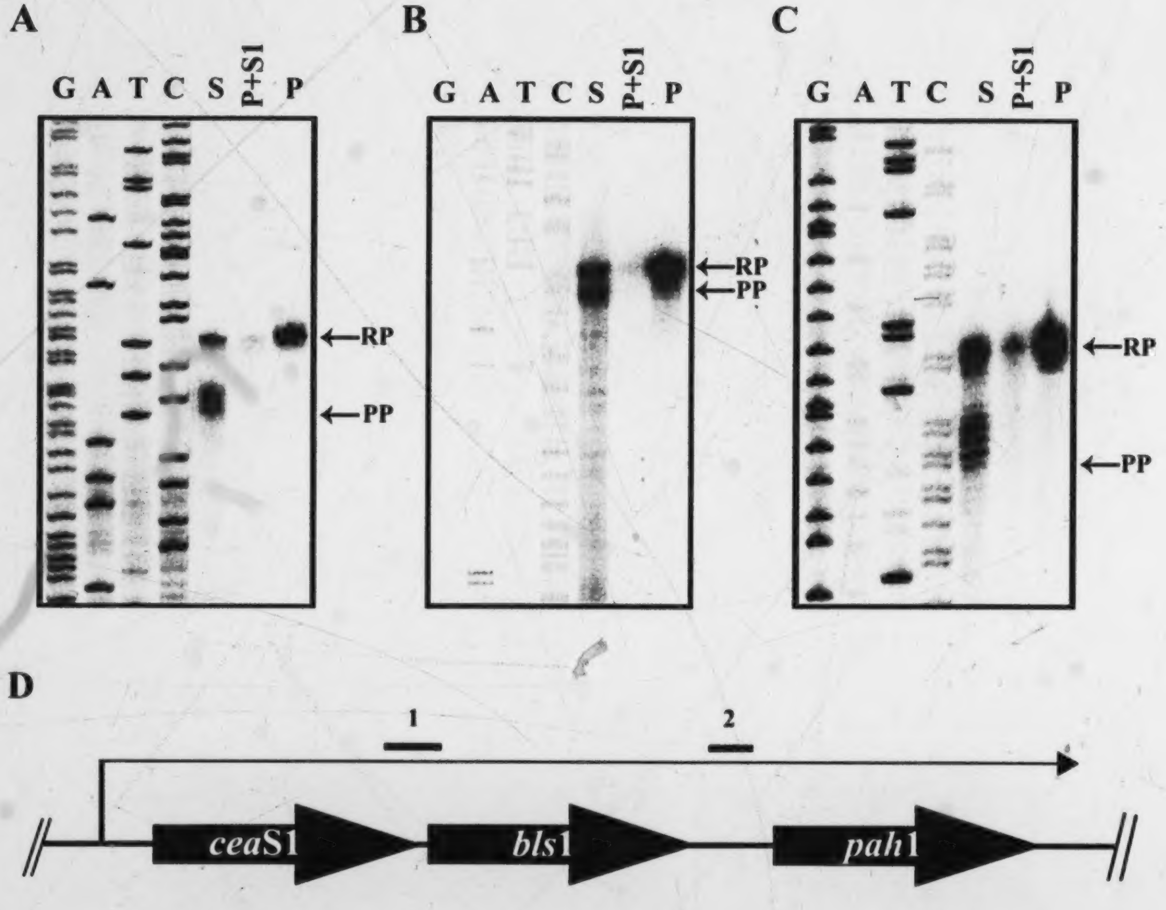
CcaR belongs to a family of transcriptional regulators called the *Streptomyces* antibiotic regulatory proteins (SARPs), which bind to specific heptameric repeats and promote transcription (Wietzorrek and Bibb, 1997)(Section 1.6). Imperfect heptameric repeats were identified in the region upstream of *cea*S2 (Figure 3.7.3) consistent with the idea that CcaR may bind directly to the *cea*S2 promoter region to regulate transcription. Such a notion however, has not yet been demonstrated experimentally. In addition, sequence analysis did not predict the presence of similar repeats upstream of *cea*S1 (Figure 3.7.3). Note: portions of this section have been published as part of the manuscript by Tahlan *et al*. (2004a).

3.8 Mapping of the *bls*1 and *pah*1 transcripts

In the clavulanic acid gene cluster, *bls*2 and *pah*2 are transcribed as part of a larger polycistronic transcript that also encodes *cea*S2 and *cas*2 (Figure 1.5)(Paradkar and Jensen, 1995). The regions upstream of *bls*1 and *pah*1 from the paralogue gene cluster were examined using S1 nuclease protection assays to determine if they were also expressed as part of a polycistronic message, along with the upstream *cea*S1 gene (Figure 3.8.1). The *bls*1 transcript was mapped using RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 hours together with a 163 bp probe extending from -137 to +16 bp relative to the proposed *bls*1 ATG start codon, which was prepared using the primers bls1-S1-For and bls1-S1-Rev along with p5.7 as template. Only full length protection of the probe was observed indicating that there was no individual TSP located in the 23 bp intergenic region between *cea*S1 and *bls*1 (Figure 3.8.1A).

The intergenic region between *bls*1 and *pah*1 was examined by S1 nuclease protection assays using a 288 bp probe extending from -257 to +21 bp relative to the proposed *pah*1 ATG codon, which was prepared using the primers pah1-S1-For and pah1-S1-Rev and the plasmid p5.7 as template. Once again, only full length protection of the probe was observed (Figure 3.8.1B). Because the DNA sequence ladder was unclear in the region of the full length-protected probe, a second 116 bp probe (-257 to -151 bp upstream of the proposed *pah*1 start codon) was also used in an S1 nuclease protection assay. Again, only full length protection of the probe was observed (Figure 3.8.1C), indicating that there was no promoter immediately upstream of *pah*1. Since the probes

Figure 3.8.1 Mapping of the polycistronic transcript encoding *bls*1 and *pah*1 using S1 nuclease protection analysis. (A) A 163 bp probe (-137 to +16 bp relative to the proposed *bls*1 start codon, plus 10 bp of non-homologous sequence) was used to map the *bls*1 transcript. (B) A 288 bp probe (-257 to +21 bp relative to the proposed *pah*1 ATG codon plus 10 bp of non-homologous sequence) was used in an attempt to map the *pah*1 transcript. (C) A 116 bp probe (-257 to -151 bp upstream of the proposed *pah*1 start codon, plus 10 bp of non-homologous sequence) was used to map the *pah*1 transcript. (A and B) The DNA sequencing ladders were prepared for size estimation using the reverse primers and template plasmids used for probe preparation. Lanes G, A, T and C is the sequencing ladder; lane S, RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 hours and subjected to analysis; lane P+S1, control lane with unprotected probe digested with S1 nuclease; lane P, undigested probe control. Bands observed due to probe/probe reannealing (RP) and full-length protection of the probes (PP), minus the 3′ non-homologous 10-nucleotide sequences, are indicated. (D) Diagrammatic representation of the *cea*S1, *bls*1 and *pah*1 polycistronic transcript. The filled arrows represent the genes, with the arrowheads indicating the direction of transcription. The fine line represents the rest of the *S. clavuligerus* chromosome. The fine arrow represents the mRNA transcript and the bars 1 and 2 represent the S1 probes used to map the *bls*1 and *pah*1 transcripts as described in A and C, respectively (diagram is not to scale).


A
G A T C S P+S1 P
←RP
←PP
B
G A T C S P+S1 P
←RP
←PP
C
G A T C S P+S1 P
←RP
←PP
D
1
2
ceaS1
bls1
pah1

used for the S1 nuclease protection studies did not cover the entire 317 bp intergenic region between *bls*1 and *pah*1, it was still possible that a TSP might be found further upstream of *pah*1. Northern analysis of RNA isolated from wild type *S. clavuligerus* grown on soy medium for 96 and 120 hours was conducted to investigate this possibility. The 288 bp probe prepared for the first *pah*1-S1 nuclease protection assay was used as the *pah*1-specific northern probe. Only a single large band of ~4.9 kb hybridized to the probe (Figure 3.8.2), indicating that *pah*1 does not have an individual promoter and that it is transcribed as part of a tricistronic operon together with *bls*1 and *cea*S1. This 4.9 kb transcript is postulated to encode *cea*S1, *bls*1 and *pah*1, as the predicted length of a transcript extending from the *cea*S1 TSP to the stop codon of *pah*1 would be approximately 4.7 kb. The next gene downstream of *pah*1 is *oat*1, which is orientated in the direction opposite to *pah*1 transcription (Figure 3.2.4), and therefore its presence on the 4.9 kb polycistronic transcript can be ruled out. In addition, computational analysis of the 126 bp intergenic region between *pah*1 and *oat*1 predicted the presence of considerable secondary structure, consisting of multiple stem loops with a cumulative ΔG ranging from −100.4 to −97.1, which could function as a transcriptional terminator (Figure 3.8.3). Note: portions of this section have been published as part of the manuscript by Tahlan *et al*. (2004a).

Figure 3.8.2 Northern analysis of wild type *S. clavuligerus* RNA using a *pah*1 specific probe. Lane 1, molecular weight marker; lanes 2 and 3, RNA from wild type *S. clavuligerus* grown on soy medium for 96 and 120 hours respectively.

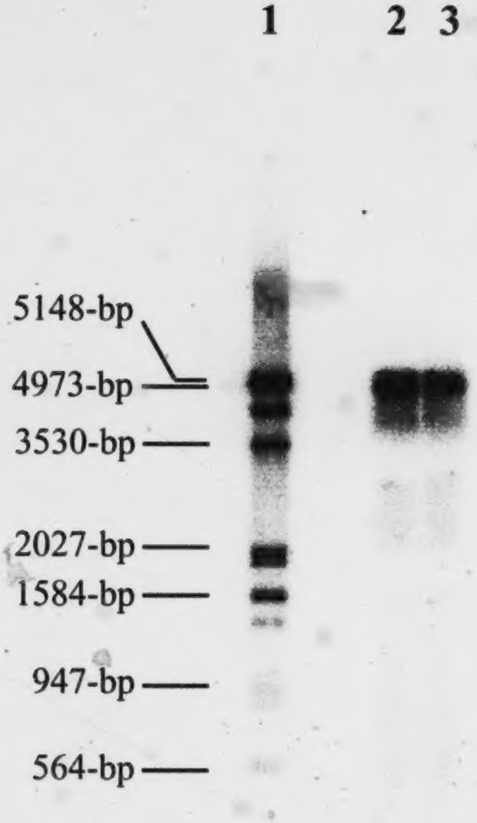
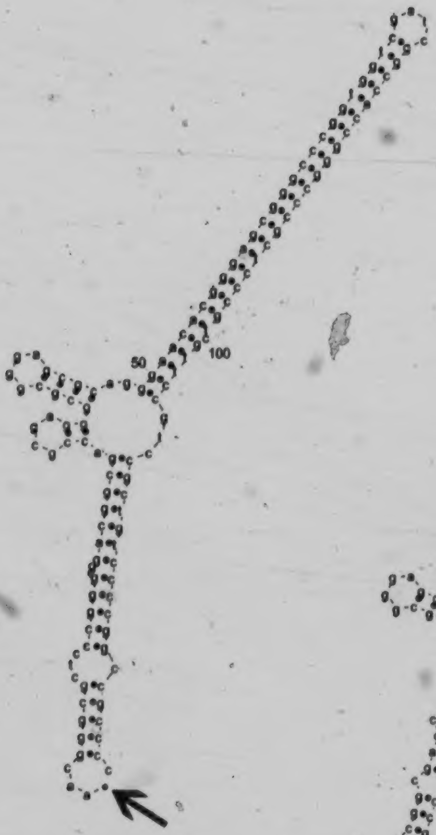
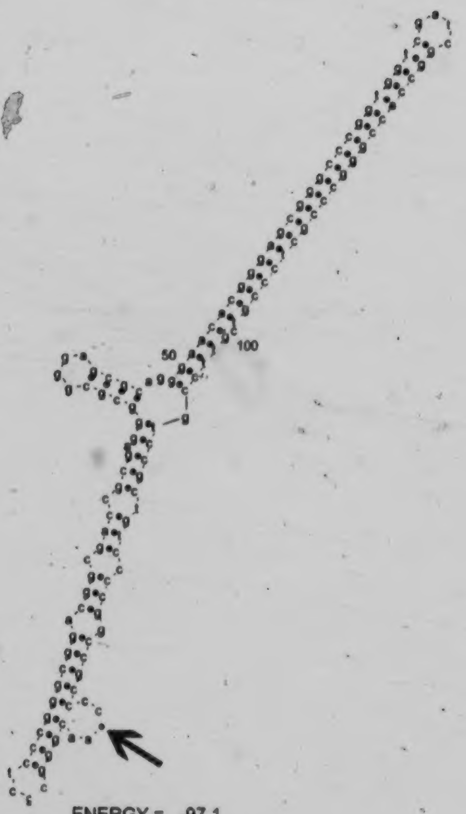

1
2
3
5148-bp
4973-bp
3530-bp
2027-bp
1584-bp
947-bp
564-bp

Figure 3.8.3 Predicted RNA secondary structure elements in the region downstream of *pah*1. The secondary structure in the 126 bp intergenic region between *pah*1 and *oat*1 was predicted and the calculated ΔG values at 28°C are shown for the two respective structures. It is postulated that the stem loops found in this region may function as transcriptional terminators. The arrow marks the first nucleotide of the sequence subjected to analysis.



ENERGY = -100.4



ENERGY = -97.1

CHAPTER 4: DISCUSSION

4. Discussion

The main objective of this thesis was the isolation and characterization of the *cea*S paralogue in *S. clavuligerus*. The regulation of the "early" genes involved in clavulanic acid and 5*S* clavam biosynthesis, and the physical linkage between the gene clusters involved in the biosynthesis of the respective metabolites, were also investigated. Results described in this thesis are compiled with results from other studies and are discussed to explain the major findings of the work presented herein.

4.1 Isolation of *cea*S1 from the paralogue gene cluster and linkage studies

Using Southern analysis, we were able to isolate a second *cea*S gene encoding carboxyethylarginine synthase in *S. clavuligerus* (Figure 1.4.1). This newly isolated gene was called *cea*S1 and the predicted amino acid sequence of the *cea*S1 gene product is 66% identical to that of CeaS2, which is encoded by the *cea*S2 gene from the clavulanic acid gene cluster (Figure 1.4.2). In addition, CeaS1 is predicted to contain all the motifs required for carboxyethylarginine synthase activity, suggesting that this putative protein also plays a similar role in clavulanic acid and 5*S* clavam metabolite biosynthesis (Section 3.2.2). The *cea*S1 gene is located near the previously isolated *pah*1 gene, which encodes a second copy of PAH in *S. clavuligerus* (Jensen *et al.*, 2004b)(Figure 3.2.4). Paralogues of other genes encoding β-LS and OAT were also found downstream of *cea*S1 (Figure 3.2.4), and were called *bls*1 and *oat*1, respectively (Jensen *et al.*, 2004b; Tahlan *et al.*, 2004c). In addition, Jensen *et al.* (Unpublished) identified three additional ORFs, called *c6p*, *c7p* and *skn* that are located downstream of *oat*1, and are predicted to encode proteins similar to aminotransferases, transcriptional regulators and two component sensor kinases, respectively (Figure 3.2.4).

Interestingly, *c6p* and *c7p* from the paralogue gene cluster are similar to *cvm*6 and *cvm*7 from the *S. clavuligerus* clavam gene cluster (Figure 1.4.3), which also encode a putative aminotransferase and a transcriptional regulator, respectively (Table 4.1). Due to this reason the gene cluster encompassing *cea*S1→*skn* was called the paralogue gene cluster as it contains a second set of the *cea*S, *bls*, *pah* and *oat* genes which are also present in the clavulanic acid gene cluster, and it contains genes that are similar to *cvm*6 and *cvm*7 from the clavam gene cluster. These findings extended the work of Jensen *et al.*

Table 4.1: Paralogous or similar genes located in the clavulanic acid, the clavam and the paralogue gene clusters of *S. clavuligerus*.

Gene pair	% Identity (Protein Level)	Known/Proposed Functions[a]	Respective Locations (Gene Clusters)[b]	Source/Reference
*cea*S1-*cea*S2	66	Carboxyethylarginine synthase	Paralogue-Clavulanic	Jensen *et al.* (2000), Perez-Redondo *et al.* (1999) and Tahlan *et al.* (2004c)
*bls*1-*bls*2	49	β-Lactam synthetase	Paralogue-Clavulanic	Bachmann et al. (1998), Jensen *et al.* (2000), McNaughton *et al.* (1998) and Tahlan *et al.* (2004c)
*pah*1-*pah*2	72	Proclaviminate amidinohydrolase	Paralogue-Clavulanic	Aidoo *et al.* (1994), Jensen *et al.* (2000), Jensen *et al.* (2004b) and Wu *et al.* (1995)
*cas*1-*cas*2	82	Clavaminate synthase	Clavam-Clavulanic	Marsh *et al.* (1992), Mosher *et al.* (1999) and Salowe *et al.* (1990)
*oat*1-*oat*2	47	Ornithine transacetylase	Paralogue-Clavulanic	Jensen *et al.* (2000) and Tahlan *et al.* (2004c)
cvm6-c6p	68	Acetylornithine aminotransferases	Clavam-Paralogue	Jensen *et al.* (Unpublished)
cvm7-c7p	33	Transcriptional Regulator	Clavam-Paralogue	Jensen *et al.* (Unpublished)

[a] Similarities at the amino acid level were determined by searching the database using the BLASTp online program.
[b] The gene clusters in *S. clavuligerus* containing the respective genes are shown

(2000), who predicted the existence of the *cea*S1, *bls*1, *pah*1 and *oat*1 genes, due to similarities in the phenotypes of the *cea*S2, *bls*2, *pah*2 and *oat*2 mutants to the *cas*2 mutant (Table 1.2), as it was already known that the second *cas*1 gene was present in *S. clavuligerus* (Section 1.4).

Results reported in this study show that the paralogue gene cluster is not located in the vicinity of the clavulanic acid gene cluster as thought initially (Section 3.5). Furthermore, the clavulanic acid, the clavam and the paralogue gene clusters are not physically linked, which is surprising as the genes involved in the biosynthesis of families of antibiotics are invariably linked in *Streptomyces* spp. to form gene clusters, which can sometimes be very large (Chater and Bibb, 1997). Due to the partial sharing of the biosynthetic pathway leading to clavulanic acid and the 5*S* clavams (Egan *et al.*, 1997), we initially expected that all of the gene clusters involved in their biosynthesis would be linked to form a still larger β-lactam supercluster in *S. clavuligerus*.

The arrangement of the respective genes in the paralogue and the clavulanic acid gene clusters is very similar (Figures 3.2.4 and 1.4.2), although the following differences were observed. In the clavulanic acid gene cluster the next gene downstream of *pah*2 is *cas*2 (Jensen *et al.*, 2000)(Figure 1.4.2), whereas the *cas*1 gene is not located downstream of *pah*1 in the paralogue gene cluster (Jensen *et al.*, 2004b; Tahlan *et al.*, 2004c), but is instead located elsewhere on the chromosome as described by Mosher *et al.* (1999). In addition, the orientation of *oat*1 in the paralogue gene cluster is inverted as compared to *oat*2 in the clavulanic acid gene cluster (Figures 3.2.4 and 1.4.2). After the *oat* genes, there is no more similarity between the respective gene clusters, as *c6p* is the next gene after *oat*1 in the paralogue gene cluster (Figure 3.2.4), whereas *opp*A1 is located downstream of *oat*2 in the clavulanic acid gene cluster (Figure 1.4.2). In addition, *orf*A is found upstream of *cea*S1 in the paralogue gene cluster (Figure 3.2.4), whereas the genes for cephamycin C biosynthesis are located upstream of *cea*S2 in the clavulanic acid gene cluster (Figure 1.4.2).

In addition to *oat*1 and *oat*2 (Section 1.4), the *arg*J gene, which encodes another ornithine acetyltransferase enzyme, is found within the *S. clavuligerus* arginine biosynthetic gene cluster (Rodriguez-Garcia *et al.*, 2000). All of the known or putative OATs analyzed to date contain the autoproteolytic cleavage motif KGXGMXXPX-

(M/L)AT(M/L)L, with cleavage taking place between the alanine and threonine residues to give the mature, active OAT holoenzyme (Abadjieva *et al.*, 2000). The motifs ^{169}KGVGMLEPDMATLL^{183} and ^{168}KGAGMLAPGLATTL^{180} are also present in OAT2 and ArgJ respectively, and proteolytic cleavage has been confirmed experimentally (de la Fuente *et al.*, 2004; Kershaw *et al.*, 2002). The predicted amino acid sequence of OAT1 in this region is ^{167}KGPGTGPAEQDDRSTLL^{183} (Tahlan *et al.*, 2004c), which deviates from the consensus sequence, although it has been suggested that OAT1 may still function as an active OAT enzyme (Elkins *et al.*, 2004). The presence of the *c6p* gene next to *oat*1 in the paralogue gene cluster is intriguing, as both *c6p* and *cvm*6 (from the clavam gene cluster) are similar to *arg*D (Jensen *et al.* Unpublished), which encodes an acetylornithine aminotransferases (AOAT) involved in arginine biosynthesis (Cunin *et al.*, 1986). ArgD catalyzes the formation of *N*-acetylornithine, which is then converted to ornithine by the action of OAT (ArgJ), thereby recycling the acetyl group in arginine biosynthesis (Cunin *et al.*, 1986). Therefore, it is plausible that *oat*1-*c6p* may be involved in providing arginine for clavulanic acid and 5*S* clavam metabolite biosynthesis, although their involvement in some other still unknown acetylation and amination reaction cannot be ruled out completely.

The *cvm*7 and *c7p* genes, which encode putative transcriptional regulators located in the clavam and paralogue gene clusters, respectively, show the lowest level of similarity (Table 4.1). Although both the *cvm*7 and the *c7p* gene products are similar to the recently described pimaricin regulator (PimR) from *S. natalensis* (Anton *et al.*, 2004), it is still not clear if they have similar functions in *S. clavuligerus* (Section 4.3). The last gene from the paralogue gene cluster to be isolated by Jensen *et al.* (Unpublished) was *skn*, which encodes a putative sensor kinase protein (Section 3.4.1). Since there is no sensor kinase gene present in the known clavulanic acid and clavam gene clusters (Figures 1.4.2 and 1.4.3), this indicates that there is no *skn* paralogue, which is located in those gene clusters.

Therefore in *S. clavuligerus*, two sets of genes located in three separate gene clusters encode the enzymes involved in the early stages of the pathway that is shared between clavulanic acid and 5*S* clavams (Table 4.1). Although exact roles of *oat*1/*oat*2 and *c6p*/*cvm*6 in clavulanic acid and 5*S* clavam biosynthesis are not known, they have been implicated in the biosynthesis of the clavam metabolites due to their location in the respective gene clusters.

4.2 Phenotypes of the *S. clavuligerus cea*S mutants

When the *cea*S2*::apr* (4B) mutant prepared by Jensen *et al.* (2000) was first analyzed, it was initially believed not to produce either clavulanic acid or the 5*S* clavams and therefore to be an exception to the rule that early genes must have paralogues. However, on repeated cultivation, on a few occasions when cell growth was especially heavy, low levels of both clavulanic acid and 5*S* clavams were detected in culture supernatants from the same mutant (Table 1.2). At that stage, it was already known that CeaS catalyzes a critical reaction, the first step in the shared clavulanic acid and 5*S* clavam biosynthetic pathways. Therefore it was presumed that a *cea*S paralogue must exist, to account for the occasional clavam metabolite production observed, but for some unknown reasons, metabolite production by the *cea*S2*::apr* mutant was very low. Since the early genes from the clavulanic acid gene cluster are expressed as a 5.3 kb polycistronic transcript (Figure 1.4.2), it was thought that the insertion of the *apr* cassette in the direction opposite to *cea*S2 transcription in the *cea*S2*::apr* mutant could lead to polar effects, thereby affecting the transcription of the genes downstream of *cea*S2 (Section 1.4), which might explain part of this low production phenotype. An earlier study conducted by Paradkar and Jensen (1995) provided evidence to support this hypothesis, as it was shown that the insertion of an *apr* cassette in the direction opposite to *cas*2 transcription abolished the transcription of the DNA sequence, immediately downstream of the insertion site. Therefore, to alleviate potential polar effects on the expression of the genes downstream of *cea*S2 in the *cea*S2*::apr* mutant, a new *cea*S2 mutant was prepared in which the *apr* disruption cassette was replaced by a frame shift mutation (*cea*S2-Fs, Section 3.1.1). This new *cea*S2-Fs mutant also produced clavulanic acid and the 5*S* clavams only rarely, but the frequency at which these metabolites were produced increased, as compared to the *cea*S2*::apr* mutant. LC-MS analysis showed that the *cea*S2-Fs mutants normally produced clavulanic acid at very low levels, making them undetectable by conventional HPLC. In addition, complementation of the *cea*S2-Fs mutant with the wild type *cea*S2 gene did not yield any conclusive results; out of two complemented *cea*S2-Fs (2-8) mutants, only one produced trace amounts of clavulanic acid. This suggested that some polar effects were still associated with the *cea*S2-Fs mutation. At the time these studies were carried out, the existence of paralogous genes (other than *cas*1) had not yet been confirmed, and the subsequent isolation of *cea*S1 and the other genes from the paralogue gene cluster made it harder to explain the polar phenotype of the *cea*S2 mutants. The presence of *cea*S1 and the other paralogues genes in these mutants should allow clavam metabolite production to occur at normal levels if these genes function as true paralogues that have similar roles in *S. clavuligerus*. In the work reported here it was shown that *cea*S1 is not a silent gene as it is transcribed in soy medium. Furthermore, in the *cca*R mutant grown on soy medium, *cea*S2 expression is reduced to near zero, yet the 5*S* clavams are still produced (Section 3.6). Presumably, there is no CeaS2 protein present in the *cca*R mutant, and therefore the production of the 5*S* clavams in this strain can be attributed to *cea*S1 expression (Section 4.4). In addition, the production of both clavulanic acid and the 5*S* clavams was reduced in the Δ*cea*S1*::apr* mutants, although not to the same degree as *cea*S2 mutants. This indicates that *cea*S1 is involved in clavam biosynthesis in *S. clavuligerus*, and therefore the phenotypes of the *cea*S2*::apr* and the *cea*S2-Fs mutants could not be explained, solely based on polar effects.

The CeaS2 protein is similar to the large subunits of the acetohydroxy acid synthases (AHAS, Section 3.1.2), which are oligomeric proteins (Eoyang and Silverman, 1984; Hill *et al.*, 1997; Vyazmensky *et al.*, 1996). In addition, the crystal structure of CeaS2 has been solved, showing that the holoenzyme is a tetramer (Caines *et al.*, 2004), and its structure is very similar to the yeast pyruvate decarboxylase (PDC)(Arjunan *et al.*, 1996). Due to the high level of similarity between the CeaS2 and the predicted CeaS1 proteins, CeaS1 is hypothesized to adopt the same three-dimensional conformation as CeaS2 (Section 3.2.2). Therefore, it was suggested that the phenotype of the *cea*S2*::apr* and *cea*S2-Fs mutants might represent a form of negative dominance, where a mutant form of the CeaS2 protein may bind to the wild type CeaS1 protein, and inhibit its activity. The fact that the corresponding effect was not observed in the Δ*cea*S1*::apr* mutants could be explained because in those mutants the entire *cea*S1 gene was deleted, eliminating the possibility of truncated CeaS1 protein expression (Section 3.2.2). To examine this hypothesis, two new *cea*S2 mutants (Δ*cea*S2*::apr* and Δ*cea*S2) were prepared in which the entire *cea*S2 gene was deleted, therefore eliminating the possibility of the formation of mutant, truncated CeaS2 protein (Section 3.1.2 and 3.1.3). In addition, the Δ*cea*S2 mutant had an in-frame deletion replacing the *cea*S2 coding sequence, and therefore should have the least effect on the expression of the downstream genes (Section 3.1.3). On analysis, both the Δ*cea*S2*::apr* and the Δ*cea*S2 mutants were still found to be very severely compromised in their ability to produce the clavam metabolites, indicating that negative dominance cannot be the explanation for the low production phenotype associated with the *cea*S2 mutants prepared earlier. The non-producing phenotype of the Δ*cea*S2 mutant was surprising and complementation of the Δ*cea*S2 mutant with wild type *cea*S2 still did not restore clavam metabolite production. This presumably indicated that some polar effect still existed in the Δ*cea*S2 mutant in addition to some other factor(s), due to which the *cea*S1→*oat*1 paralogues were unable to restore clavam metabolite biosynthesis in the *cea*S2 mutants.

In the case of the *lat*, *pcb*AB and *pcb*C genes from the cephamycin gene cluster of *S. clavuligerus* (Figure 1.3.2), it was shown that a *lat* mutation prepared by insertional inactivation, not only abolished the transcription of *lat*, *pcb*AB and *pcb*C from the *lat* promoter, but also severely affected the expression of *pcb*C from its own promoter (Alexander *et al.*, 2000). The *lat* and *pcb*C promoters are separated by approximately 13 kb of DNA in *S. clavuligerus*, and it was proposed that *pcb*C promoter activity was dependent on transient changes in the local DNA topology of the *pcb*C promoter region, due to transcription of the upstream genes (Alexander *et al.*, 2000). The transcriptional arrangement of the "early" genes from the clavulanic acid gene cluster is complex and is still not fully understood. In addition, some uncharacterized transcripts spanning the *cas*2-*oat*2 and *oat*2-*opp*A1 intergenic regions are thought to exist (Section 1.4). Previous studies have shown that the *oat*2 mutant does not produce the 5*S* clavams (except for trace levels of alanylclavam) and the *opp*A1 mutant is completely blocked in clavulanic acid biosynthesis (Jensen *et al.*, 2000; Lorenzana *et al.*, 2004). Therefore, if polar effects in the *cea*S2 mutants can knock out the transcription of *bls*2 and *pah*2, and inhibit *cas*2, *oat*2 and *opp*A1 transcription, this would lead to the loss of clavulanic acid and 5*S* clavam production, even though an intact paralogue complement of genes is present in these mutants. The main evidence to back this hypothesis comes from a study conducted by Jensen *et al.* (Unpublished), where it was shown that a *S. clavuligerus* mutant generated by random mutagenesis and thought to be defective in either *cea*S2 or *bls*2 was best complemented (up to 30% of wild type level) when a DNA fragment encompassing *cea*S2→*opp*A1 was used.

In another study, it was shown that the transcription/translation of *pcb*C in *E. coli* from an upstream promoter was dependent on the translation of the upstream *pcb*AB gene (Doran *et al.*, 1990b). Therefore, it is possible that premature transcriptional termination in the *cea*S2 mutants due to uncoupled transcription-translation could eliminate the transcription of the downstream genes or affect mRNA stability. A similar phenomenon may still occur in the Δ*cea*S2 mutant despite the in-frame nature of the deletion, as the "scar" sequence left after the deletion has an average G+C content of 48%, which is very different from the average G+C content of *Streptomyces* DNA, and could therefore effect translation (Bibb *et al.*, 1984). Although there are no *bld*A-dependent TTA codons in the "scar" sequence, there is a CTA codon, which after TTA and TAA is the third most rare codon found in *Streptomyces* coding sequences. Therefore, if stalling or slowing down of the ribosome at rarely used codons could uncouple transcription-translation, this could lead to premature transcriptional termination (Gowrishankar and Harinarayanan, 2004) or the cleavage of the downstream-untranslated mRNA (Sunohara *et al.*, 2004). In addition, the *cea*S2 mutants are predicted to be highly polar as an intergenic region of only 2 bp separates *cea*S2 from the downstream *bls*2 gene (Figure 3.1.15), and there is no evidence of a potential ribosome-binding site immediately upstream of *bls*2. Therefore, the translation of both *cea*S2 and *bls*2 could be linked, although this has not been verified experimentally. This would explain the low-level clavam production observed in the *cea*S2 mutants by LC-MS. As well, occasional transcription/translation of the genes downstream of *cea*S2 could lead to the rare-production of clavulanic acid and the 5*S* clavams, which is detected by conventional HPLC analysis.

Another explanation for the severe phenotype of the *cea*S2 mutants could be that under normal conditions, *cea*S1 is not functionally equivalent to *cea*S2. If the *cea*S1 and *cea*S2 gene products have different substrate specificities, then complementation of the *cea*S2 mutants should restore clavam production, which was not observed. Therefore, the notion that *cea*S1 and *cea*S2 have different roles in clavam metabolite biosynthesis still cannot explain the phenotype of *cea*S2 mutants, as *bls*1→*oat*1 should be available for clavam production in these mutants, if polarity is also an issue.

In considering other explanations for the extreme phenotype of the *cea*S2 mutants, the possibility of channeling of intermediates between the different proteins involved in clavulanic acid and 5*S* clavam metabolite biosynthesis was also contemplated (Kershaw *et al.*, 2002). Perhaps the correlation between relative arrangements of the genes from the clavulanic acid and the paralogue gene clusters (Figures 1.4.2 and 3.2.4) and the order in which the encoded proteins catalyze reactions in the biosynthesis of the clavams (Figure 1.4.1), may indicate that the respective proteins must associate in an orderly fashion (S.E. Jensen, Personal communication). This could explain why complementation of the *cea*S2 mutants *in trans* did not restore clavam metabolite production. Furthermore, complementation observed due to the DNA fragment encompassing *cea*S2→*opp*A1 may indicate that CeaS2 needs to associate with all the proteins encoded on that fragment in a specific order for clavam metabolite production to take place. However for this to be true it would also require that CeaS1 not be involved in clavulanic acid and 5*S* clavam metabolite biosynthesis under normal growth conditions. As well, such a protein association and channeling mechanism cannot be predicted for the *bls*1/ *bls*2, *pah*1/ *pah*2, *cas*1/*cas*2 and *oat*1/*oat*2 paralogues, as mutants defective in either copies of these genes still produce clavam metabolites (Jensen *et al.*, 2000; Jensen *et al.*, 2004b; Tahlan *et al.*, 2004c). Although protein-protein interactions with the channeling intermediates in the biosynthesis of the clavams present an attractive hypothesis, it needs to be viewed with caution, as there is no evidence to back it up experimentally.

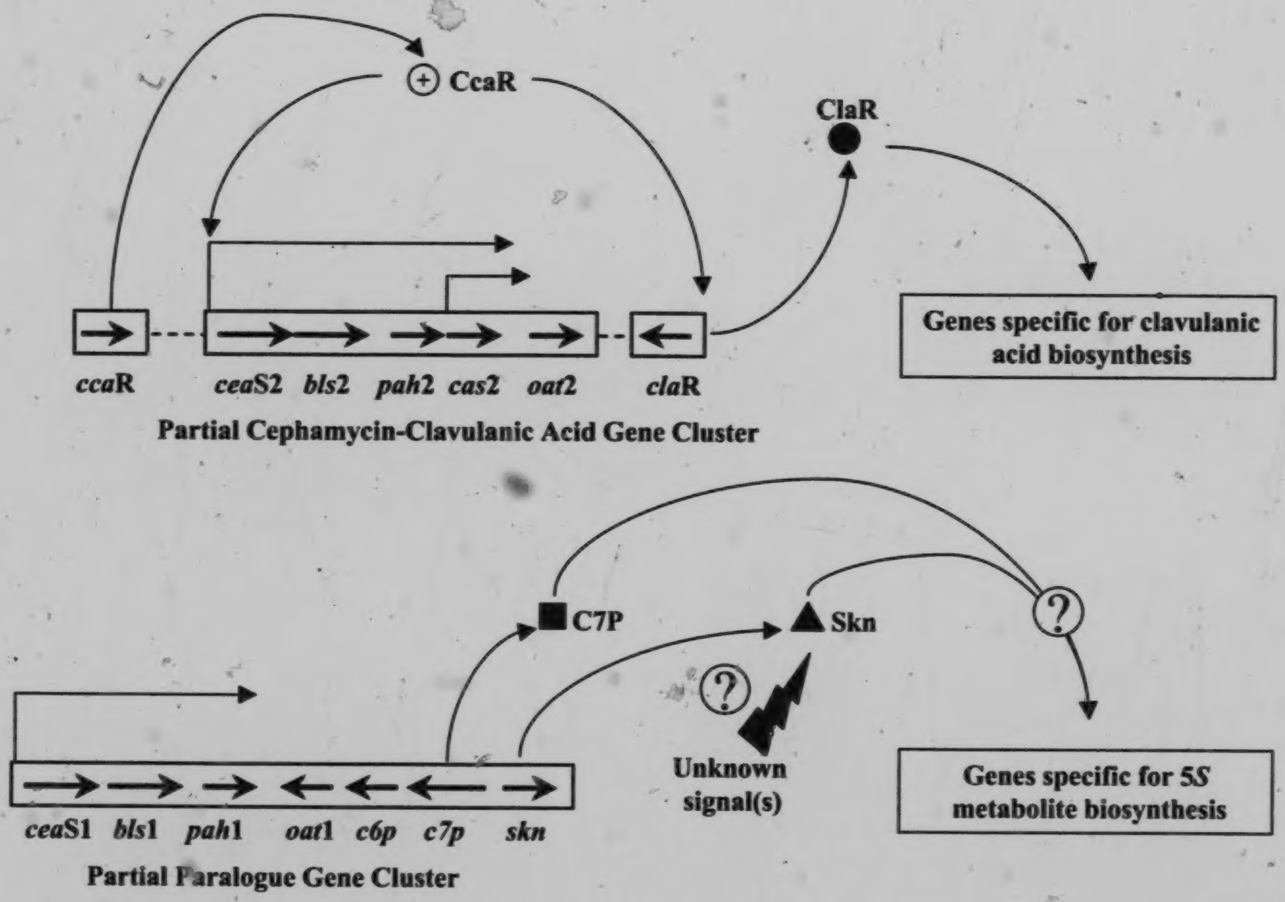
Lastly, LC-MS analysis showed that the low level of clavulanic acid production observed in the *cea*S2-Fs mutants, was completely abolished in the Δ*cea*S1*::apr/cea*S2-Fs double mutants (Section 3.3.4). This reiterates that *cea*S1 is involved in clavam metabolite biosynthesis, lending strength to the first hypothesis that the polarity of the *cea*S2 mutants may be responsible for their severe phenotypes, although the interplay of polar and additional effects cannot be ruled out completely.

4.3 Phenotypes of the *S. clavuligerus* Δ*c7p::apr* and Δ*skn::apr* mutants

The *c7p* and *skn* ORFs from the paralogue gene cluster, which are present on the end opposite to *cea*S1 (Figure 3.2.4), and encode a putative transcriptional regulator and two-component sensor kinase, respectively, were subjected to mutational analysis to determine if they are involved in clavulanic acid and/or 5*S* clavam biosynthesis. As described earlier (Sections 3.4 and 4.1), *c7p* is similar to *cvm*7 from the clavam gene cluster (Figure 1.4.3, Table 4.1). Jensen *et al.* (unpublished) prepared a *cvm*7 mutant and demonstrated that it had no apparent effect on either clavulanic acid or 5*S* clavam metabolite biosynthesis. When a *c7p* mutant was prepared in the present study, it did not produce any detectable levels of the 5*S* clavams, whereas it still produced clavulanic acid (Section 3.4.3), indicating that *c7p* is involved in regulating the biosynthesis of the 5*S* clavams only and not that of clavulanic acid. In addition, these results show that the *cvm*7 and the *c7p* gene products do not have similar functions in *S. clavuligerus* due to the differences in the phenotypes of the two mutants. Therefore, it is possible that C7P functions as the pathway-specific regulator for 5*S* clavam biosynthesis, analogous to the role of ClaR in clavulanic acid biosynthesis (Figure 4.1)(Section 1.5).

The *skn* mutants were similar in their phenotypes to the *c7p* mutants described above as they were compromised in 5*S* clavam metabolite biosynthesis, but retained the ability to produce clavulanic acid (Section 3.4.3). The predicted amino acid sequence of Skn suggests that it is located in the cytoplasm and it shows the presence of GAF domains (Section 3.4.1), which are thought to bind small ligand molecules to modulate the activity of the target protein (Aravind and Ponting, 1997). Similar sensor kinase proteins are also found in *S. coelicolor*, although the signals that they sense and the pathways that they affect are still not known (Hutchings *et al.*, 2004). Since GAF domains bind small ligands, it is possible that Skn functions as a soluble protein that binds some still unknown small molecule(s), allowing it to phosphorylate other proteins that may then function as activators or repressors, to either directly or indirectly regulate 5*S* clavam metabolite biosynthesis in *S. clavuligerus* (Figure 4.1). Since the identities of the ORFs located downstream of *skn* are not yet known (Figure 3.2.4), it will be interesting to see if there is a paired response regulator located next to *skn* once this region of the *S. clavuligerus* chromosome is sequenced.

Figure 4.1 Proposed roles of the different regulatory proteins described in this study. The partial cephamycin-clavulanic acid and paralogue gene clusters are shown with the arrows representing the genes and the arrowheads indicating the direction of transcription. The bent arrows represent the transcripts encoding the early genes from the clavulanic acid and paralogue gene clusters and the filled shapes represent the respective proteins as indicated. The open circle with the question mark indicates speculated steps that have not been verified experimentally or steps where additional proteins could be involved.


CcaR
ClaR
Genes specific for clavulanic acid biosynthesis
ccaR
ceaS2
bls2
pah2
cas2
oat2
claR
Partial Cephamycin-Clavulanic Acid Gene Cluster
C7P
Skn
?
?
Unknown signal(s)
Genes specific for 5S metabolite biosynthesis
ceaS1
bls1
pah1
oat1
c6p
c7p
skn
Partial Paralogue Gene Cluster

The phenotypes of the *c7p* and *skn* mutants show that the paralogue gene cluster also contains genes that are only involved in *5S* clavam metabolite biosynthesis and not that of clavulanic acid. In addition, it also demonstrates that the *5S* clavam-specific genes are present in the known clavam and paralogue gene clusters and not in the clavulanic acid gene cluster. The significance of this gene arrangement it not clear yet, but it may imply differential evolution and/or regulation of the genes involved in clavulanic acid and *5S* clavam metabolite biosynthesis in *S. clavuligerus* (Section 4.5).

4.4 Transcriptional mapping of the early genes from the paralogue gene cluster

A combination of S1 nuclease protection and primer extension analyses was used to map the start site(s) for the transcripts containing *cea*S1, *bls*1 and *pah*1 from the paralogue gene cluster (Section 3.7). The *cea*S1 transcript originated from a single TSP located at a C residue 98 bp upstream from the start codon, but the *cea*S1 promoter region showed little similarity to known *Streptomyces* promoters (Figure 3.7.3), a reflection of the large diversity found in *Streptomyces* promoter sequences (Strohl, 1992). When the *bls*1 and the *pah*1 transcripts were analyzed by S1 nuclease protection assays, no individual TSP was detected for either of these genes and only full-length protection of the probes was observed. This suggested that neither *bls*1 nor *pah*1 has its own dedicated promoter and that the upstream *cea*S1 promoter drives the transcription of these genes.

The large intergenic region of 317 bp that separates *bls*1 and *pah*1 was not fully covered by the probes used for S1 nuclease protection studies. Therefore, northern analysis was used to ensure that any *pah*1 transcript originating from a promoter located further upstream was not missed. Since *pah*1 and *pah*2 share 72% end to end identity at the nucleotide level (Jensen *et al.*, 2004b), the probe used in northern analysis was carefully chosen to be specific for *pah*1. The only hybridization seen was to a band approximately 4.9 kb in size, again indicating that *pah*1 was transcribed as part of a large polycistronic transcript (Figure 3.8.2), which is postulated to encode *cea*S1, *bls*1 and *pah*1 (Section 3.8).

The transcriptional arrangement of *cea*S1, *bls*1 and *pah*1 is similar to that of their paralogous counterparts from the clavulanic acid gene cluster (Figure 1.4.2). The *cea*S2, *bls*2, *pah*2 and *cas*2 genes from the clavulanic acid gene cluster are transcribed as a 5.3 kb polycistronic transcript. In addition, *cas*2 is also transcribed as a 1.2 kb monocistronic transcript from its own promoter (Figure 1.4.2, Section 1.4). The most significant difference between the 5.3 kb transcript arising from the clavulanic acid gene cluster and the 4.9 kb transcript arising from the paralogue gene cluster is the absence of the *cas*1 coding sequence in the paralogue gene cluster, which is located in the clavam gene cluster (Mosher *et al.*, 1999) and is expressed as a 1.4 kb monocistronic transcript (Paradkar and Jensen, 1995).

The transcript encoding *cea*S2 was also mapped and was also found to arise from a single transcriptional start point located 51 bp upstream of the *cea*S2 start codon. As was the case for *cea*S1, the proposed *cea*S2 promoter region did not show any significant similarity to any known *Streptomyces* promoters (Figure 3.7.3). Since S1 nuclease and primer extension analyses were used to identify all of the TSPs described in this study, it should be noted that both of the analysis methods employed are only predictive of the TSP, provided the mRNA is not processed.

4.5 Nutritional regulation and effect of CcaR on clavam metabolite biosynthesis

Since the phenotypes of mutants defective in the paralogous genes that encode CeaS, β-LS and PAH are similar to the *cas*1 and *cas*2 mutants (Section 1.4), the nutritional regulation of *cea*S1 and *cea*S2 was examined to determine if they are regulated in a manner similar to the *cas*1 and *cas*2 paralogues (Section 3.6.1.1). RT-PCR analysis of wild type *S. clavuligerus* RNA demonstrated that *cea*S1 is transcribed in soy medium only and not in SA medium (Figure 3.6.1). Under the same conditions and using the same RNA preparations, *cea*S2 is transcribed in both soy and SA media at comparable levels (Figure 3.6.1). The *cea*S1 and the *cea*S2 promoter regions were also subcloned in front of a promoterless *egfp* gene, and EGFP expression was used as a reporter to monitor transcription driven by the respective promoters (Section 3.6.1.2). Confocal microscopy was used to detect fluorescence due to EGFP expression/excitation, and the results confirmed that the *cea*S1 promoter is active in soy medium only and not in SA medium, whereas the *cea*S2 promoter is active in both media tested (Figure 3.6.3). Results obtained from confocal microscopy were confirmed by western analysis, which indicated that the fluorescence observed in the samples was due to true EGFP expression. Since *cea*S1, *bls*1 and *pah*1 are only expressed as a 4.9 kb polycistronic message, it can be inferred that *bls*1 and *pah*1 will show the same general trend of nutritional regulation as *cea*S1. In combination, the results indicate that *cea*S1, *bls*1 and *pah*1 are expressed in soy medium but not in SA medium, whereas *cea*S2, *bls*2, *pah*2 and *cas*2 are expressed in both soy and SA medium. This explains the clavulanic acid and 5*S* clavam-producing phenotypes observed when mutants defective in these genes were prepared and tested in previous studies (Jensen *et al*., 2000; Mosher *et al*., 1999; Paradkar and Jensen, 1995; Tahlan *et al*., 2004c).

The *cca*R gene from the cephamycin gene cluster encodes a pathway-specific transcriptional regulator that coordinates the production of both cephamycin C and clavulanic acid (Alexander and Jensen, 1998; Perez-Llarena *et al*., 1997b). Its effect on clavulanic acid biosynthesis is exerted, at least in part, through activation of expression of a second pathway-specific transcriptional regulator, ClaR, from the clavulanic acid gene cluster (Perez-Redondo *et al*., 1998)(Section 1.5). Previous studies had shown that *cea*S2 expression is not under the control of *cla*R (Paradkar *et al*., 1998), but the detailed effects of CcaR on *cea*S1 and *cea*S2 expression was not known.

The results of this study indicate that *cea*S1 transcription is unaffected in the *cca*R mutant as compared to the wild type strain, whereas the transcription of *cea*S2 is almost completely eliminated in the *cca*R mutant (Section 3.6.2). Therefore, CcaR controls production of clavulanic acid through at least two routes, an indirect route mediated by ClaR, and a second route, which may involve CcaR directly regulating *cea*S2 promoter activity (Figure 4.1). In addition, since *cea*S2, *bls*2 and *pah*2 are expressed as a polycistronic transcript (Paradkar and Jensen, 1995), most likely from the *cea*S2 promoter, it can be inferred that CcaR also regulates the expression of *bls*2 and *pah*2.

CcaR belongs to a family of transcriptional regulators called the *Streptomyces* antibiotic regulatory proteins (SARPs), which bind to specific heptameric repeats and promote transcription (Wietzorrek and Bibb, 1997)(Section 1.5). Imperfect heptameric repeats can be identified in the region upstream of *cea*S2 (Figure 3.7.3) consistent with the idea that CcaR may bind directly to the *cea*S2 promoter region to regulate transcription. Such a notion however, has not yet been demonstrated experimentally, and it is also possible that CcaR asserts its effect indirectly through additional proteins. Since *cla*R is not expressed in the *cca*R mutant (Perez-Redondo *et al.*, 1998), it can also be inferred that *cla*R has no effect on *cea*S1 transcription, which was unaffected in the *cca*R mutant. Similarly, due to the polycistronic nature of the transcript encoding *cea*S1, *bls*1 and *pah*1, it follows that neither *bls*1 nor *pah*1 are affected by CcaR or ClaR. Therefore, it seems that the corresponding genes from the paralogues and the clavulanic acid gene clusters are regulated differently both in terms of nutrition and the effect of CcaR on their expression.

4.6 Purpose of paralogues genes in clavam metabolite biosynthesis

The reason why *S. clavuligerus* possesses two sets of genes encoding enzymes involved in the early stages of clavulanic acid and 5*S* clavam biosynthesis is still unclear. One suggestion is that this could be a strategy to increase precursor and metabolite flux through the shared part of the pathway by increasing gene dosage. This is consistent with the observation that both sets of the paralogous genes are expressed in complex soy medium where precursor availability and growth would support greater metabolite production levels than are possible on defined SA medium where only the *cea*S2→*oat*2 set of paralogues is expressed. In addition, the increased production of these secondary metabolites may be of greater advantage in complex medium to ward off competition, especially that posed by faster growing organisms. In defined media such as SA, the expression of only one set of paralogous genes may suffice to provide an adequate supply of precursors, resulting in lower levels of clavulanic acid and 5*S* clavam production. In addition, the phenomenon of paralogous genes encoding enzymes for clavam metabolite biosynthesis in *S. clavuligerus* seems to be limited to the early part of the pathway that is shared between clavulanic acid and the 5*S* clavams (Figure 1.4.1). This is indicated by the fact that mutants defective in genes thought to be involved specifically in the biosynthesis of either clavulanic acid or the 5*S* clavams are completely blocked in the production of the respective metabolites, under the tested nutritional conditions (Jensen *et al.*, 2000). Therefore, it is possible that single copies of genes that encode the enzymes involved in this "late" portion of the pathway are sufficient to maintain the production of the respective metabolites, with no requirement for extra gene dosage.

Another explanation put forth is that the two sets of paralogous genes may belong to two separate pathways, with one leading to clavulanic acid and the other to the 5*S* clavams. Since the two pathways proceed through common early steps, a sharing of biosynthetic intermediates results, but the two pathways may be regulated differently. This is consistent with the observation that, although clavulanic acid production was knocked out in the *cca*R mutant, the 5*S* clavams were still produced when grown on soy medium. The *cca*R mutant shows *cea*S1 transcription equivalent to that seen in the wild type strain and still produces wild type levels of 5*S* clavams, whereas *cea*S2 transcription is almost absent and clavulanic acid production is lost (Section 3.6.2). This suggests that *cea*S1, *bls*1 and *pah*1 may be more closely associated with the production of the 5*S* clavams, via a CcaR independent pathway, whereas *cea*S2 and its related paralogues are associated with clavulanic acid production and are regulated by CcaR (Figure 4.1). This is an attractive hypothesis from the point of view that the producer organisms would be best served by coordinating production of a β-lactam antibiotic (cephamycin C) with production of a β-lactamase inhibitor (clavulanic acid) through the action of a common regulator (CcaR). In contrast, no apparent advantage would be gained by coordinating production of the 5*S* clavams with production of cephamycin C. Furthermore, the *c7p* and *skn* mutants indicate that the late steps leading specifically to the 5*S* clavams are differently regulated as compared to those leading to clavulanic acid, which are regulated by ClaR (Figure 4.1).

The finding that the clavulanic acid, the clavam and the paralogue gene clusters are three separate entities may suggest that two different pathways are involved in the biosynthesis of clavulanic acid and the 5*S* clavams. This raises further questions regarding the origins and evolution of the respective gene clusters. It is still not clear if the clavulanic acid, the clavam or the paralogue gene clusters arose by some sort of gene duplication event in *S. clavuligerus*, or if one or more of them were acquired later on by Horizontal Gene Transfer (HGT). It has been suggested that the clavulanic acid gene cluster arose in a cephamycin and 5*S* clavam producing *Streptomyces* by chromosomal duplication of the clavam gene cluster, followed by the acquisition of the late reactions specific for clavulanic acid biosynthesis (Challis and Hopwood, 2003). As well, the linkage of the cephamycin and clavulanic acid gene clusters and the finding that all

Streptomyces spp. that produce clavulanic acid (Table 1.1) also produce cephamycins, suggests that both the ceph-clav gene clusters are inherited together. Therefore if the clavulanic acid gene cluster was gained by *S. clavuligerus* later on in its evolutionary history by HGT, then the correct term to define the two sets of genes encoding the enzymes involved in the early stages of clavulanic acid and 5*S* clavam metabolite biosynthesis would be xenologus, not paralogous (Gevers *et al.*, 2004).

The functional holoenzyme forms of CeaS2 (Caines *et al.*, 2004), β-LS2 (Miller *et al.*, 2001), PAH2 (Elkins *et al.*, 2002) and OAT2 (Elkins *et al.*, 2004; Kershaw *et al.*, 2002) have all been characterized structurally and have been shown to be oligomers. Since these proteins were overexpressed and purified from *E. coli*, only homo-oligomers were observed. It is reasonable to expect that when the corresponding homo-oligomeric forms of CeaS1, BLS1 and PAH1 are expressed and purified, they may have somewhat different activities or kinetic properties, just as was seen for the CAS1 and CAS2 monomers (Salowe *et al.*, 1990). These differences in activities may be important under the specific nutritional conditions when each of these isoenzymes is expressed. It is also possible that within *S. clavuligerus* the two isozymic forms of each protein can form hetero-oligomers, which could provide another mechanism to modulate enzyme activity based on nutritional and precursor availability.

Overall, based on the ideas discussed above, one is tempted to speculate about the origins and the roles of paralogous genes in clavam metabolite biosynthesis in *S. clavuligerus*, but additional work is required before any firm conclusions can be drawn.

4.7 Conclusion and proposed future research

The main objective of this thesis, which was the isolation and characterization of the *cea*S paralogue in *S. clavuligerus*, was fulfilled. It was also shown that the genes involved in the biosynthesis of the clavam metabolites are located in three separate gene clusters in *S. clavuligerus*. To our knowledge, this is the first example to show that three unlinked clusters of genes are involved in the production of a family of secondary metabolites that arise from a shared biosynthetic pathway in *Streptomyces*. Since neither the clavulanic acid nor the 5*S* clavam pathways have been fully elucidated as-of-yet, further studies will be required before the relationships between the different metabolites and the evolution of the gene clusters involved in their biosynthesis can be understood. Lastly, it was demonstrated that the paralogous pairs of genes encoding the enzymes involved in the early stages of clavam metabolite biosynthesis are differently regulated in *S. clavuligerus*.

The work described in this thesis raises further questions, which are potential topics of future research. Future work should entail a closer examination of the severe phenotype of the *cea*S2 mutants, which would involve complementation of the *cea*S2 mutants with a DNA fragment encompassing *cea*S2→*opp*A1 (Figure 1.4.2). In the complementation studies conducted by Jensen *et al.* (Unpublished), which used the same DNA fragment, the exact gene in which the mutation resided was not known, although it was thought to lie in either *cea*S2 or *bls*2. Therefore, complementation of the *cea*S2 mutants should confirm whether this DNA fragment can actually complement the *cea*S2 mutants, indicating the existence of polar effects. The expression of the genes downstream of *cea*S2 in the *cea*S2 mutants should also be analyzed, to determine if these genes are affected at the transcriptional or the translational level. Protein binding studies can be carried out to determine if CeaS associates with other proteins involved in clavam metabolite biosynthesis in *S. clavuligerus*. As well, the putative CeaS1 protein can be overexpressed, purified and characterized, similar to what was done for CeaS2 (Caines *et al.*, 2004; Khaleeli *et al.*, 1999), to determine its substrate specificity and kinetic properties. Protein binding studies should indicate if CeaS1 and CeaS2 interact to form a hetero-oligomer, and if they do, it can then be determined what the effect of this interaction on the activity of the holoenzyme is. These studies would aid in not only explaining the phenotype of the *cea*S2 mutants, but would also provide further insight into the roles of the paralogous genes in clavam metabolite biosynthesis in *S. clavuligerus*.

Although the expression and purification of soluble CcaR protein has always been very challenging, further analysis is required to determine if CcaR actually binds to the heptameric repeats identified in the *cea*S2 promoter region (Figure 3.7.3). Furthermore, if CcaR binding is observed, then detailed DNA footprinting analysis can be performed to identify the exact sequence of the binding site. Since the *c7p* and *skn* mutants did not produce the 5*S* clavams, transcriptional studies can be performed to determine the genes, which are controlled either directly or indirectly by Skn and C7P. As well, the transcription of the early genes from the paralogue gene cluster should also be examined in the *c7p* and *skn* mutants, to determine if their expression is somehow influenced by Skn or C7P. If the transcription of *cea*S1→*pah*1 does turn out to be dependent on Skn or C7P, then this would provide further evidence to suggest that the paralogue gene cluster is indeed more a part of the 5*S* clavam biosynthetic pathway, rather than the clavulanic acid pathway. In addition, the region of the *S. clavuligerus* chromosome downstream of *skn* needs to be analyzed to determine if there are additional ORFs present in this region, which are involved in either clavulanic acid or 5*S* clavam metabolite biosynthesis.

Therefore, there is ample scope for further studies involving the regulation and biosynthesis the clavam metabolites in *S. clavuligerus*. As well, many of the ideas described above are already being investigated by other workers (Jensen *et al.*, unpublished), and should lead to some interesting results in the near future.

CHAPTER 5: REFERENCES

5. References

Abadjieva, A., Hilven, P., Pauwels, K. and Crabeel, M. (2000). The Yeast *ARG7* gene product is autoproteolized to two subunit peptides, yielding active ornithine acetyltransferase. *J. Biol. Chem.*, **275**, 11361-11367.

Abraham, E.P., Newton, G.G.F., Schenck, J.R., Hargie, M.P., Olson, B.H., Schuurmans, D.M., Fisher, M.W. and Fusari, S.A. (1955). Identity of cephalosporin N and synnematin B. *Nature*, **176**, 551.

Aguilar, A. and Hopwood, D.A. (1982). Determination of methylenomycin A synthesis by the pSV1 plasmid from *Streptomyces violaceus-ruber* SANK 95570. *J. Gen. Microbiol.*, **128**, 1893-901.

Aharonowitz, Y., Bergmeyer, J., Cantoral, J.M., Cohen, G., Demain, A.L., Fink, U., Kinghorn, J., Kleinkauf, H., MacCabe, A., Palissa, H., Pfeifer, E., Schwecke, T., van Liempt, H., von Dohren, H., Wofe, S. and Zhang, J. (1993). δ-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase, the multienzyme integrating the four primary reactions in beta-lactam biosynthesis, as a model peptide synthetase. *Biotechnol.*, **7**, 807-10.

Aharonowitz, Y. and Demain, A.L. (1978). Carbon catabolite regulation of cephamycin production in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.*, **14**, 159-164.

Aidoo, K.A., Paradkar, A.S., Alexander, D.C. and Jensen, S.E. (1993). Use of recombinant DNA techniques to study the production of β-lactam compounds in *Streptomyces clavuligerus*. In Gullo, V.P., Hunter-Cevera, J.C., Cooper, R. and Johnson, R.K. (eds.), *Developments in Industrial Microbiology series*. Society for Industrial Microbiology, Fredricksburg, VA, Vol. 33.

Aidoo, K.A., Wong, A., Alexander, D.C., Rittammer, R.A. and Jensen, S.E. (1994). Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. *Gene*, **147**, 41-6.

Alexander, D.C., Brumlik, M.J., Lee, L. and Jensen, S.E. (2000). Early cephamycin biosynthetic genes are expressed from a polycistronic transcript in *Streptomyces clavuligerus*. *J. Bacteriol.*, **182**, 348-56.

Alexander, D.C. and Jensen, S.E. (1998). Investigation of the *Streptomyces clavuligerus* cephamycin C gene cluster and its regulation by the CcaR protein. *J Bacteriol*, **180**, 4068-79.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403-410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389-3402.

Alvarez, E., Cantoral, J.M., Barredo, J.L., Diez, B. and Martin, J.F. (1987). Purification to homogeneity and characterization of acyl coenzyme A:6-aminopenicillanic acid acyltransferase of *P. crysogenum. Antimicrob. Agents Chemother.*, **31**, 1675-1682.

Ambler, R.P. (1980). The structure of beta-lactamases. *Philos. Trans. R. Soc. London, Ser. B, Biological Sciences*, **289**, 321-31.

Anton, N., Mendes, M.V., Martin, J.F. and Aparicio, J.F. (2004). Identification of PimR as a positive regulator of pimaricin biosynthesis in *Streptomyces natalensis. J. Bacteriol.*, **186**, 2567-75.

Aravind, L. and Ponting, C.P. (1997). The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem. Sci.*, **22**, 458-459.

Arias, P., Fernandez-Moreno, M.A. and Malpartida, F. (1999). Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J. Bacteriol.*, **181**, 6958-68.

Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D. and Jordan, F. (1996). Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the Yeast *Saccharomyces cerevisiae* at 2.3 A° resolution. *J. Mol. Biol.*, **256**, 590-600.

Bachmann, B.O., Li, R. and Townsend, C.A. (1998). β-Lactam synthetase: a new biosynthetic enzyme. *Proc. Natl. Acad. Sci. U S A*, **95**, 9082-6.

Baggaley, K.H., Brown, A.G. and Schofield, C.J. (1997). Chemistry and biosynthesis of clavulanic acid and other clavams. *Nat. Prod. Rep.*, **14**, 309-33.

Bailey, C.R. and Winstanley, D.J. (1985). Inhibition of restriction in *Streptomyces clavuligerus* by heat treatment. *J. Gen. Microbiol.*, **132**, 2945-2947.

Baldwin, J.E., Adlington, R.M., Bryans, J.S., Bringhen, A.O., Coates, J.B., Crouch, N.P., Lloyd, M.D., Schofield, C.J., Elson, S.W., Baggaley, K.H., Cassels, R. and Nicholson, N.H. (1991). Isolation of dihydroclavaminic acid, an intermediate in the biosynthesis of clavulanic acid. *Tetrahedron*, **47**, 4089-4100.

Baldwin, J.E., Fujishima, Y., Goh, K.-C. and Schofield, C.J. (1[illegible]). Enzymes of valclavam biosynthesis. *Tetrahedron Lett.*, **35**, 2783-2786.

Baldwin, J.E., Lloyd, M.D., Wha-Son, B., Schofield, C.J., Elson, S.W., Baggaley, K.H. and Nicholson, N.H. (1993). A substrate analog study on clavaminic acid synthase: possible clues to the biosynthetic origin of proclavaminic acid. *J. Chem. Soc. Chemical. Commun.*, 500-502.

Baitz, R.H. (1980). Genetic recombination by protoplast fusion in *Streptomyces*. *Dev. Ind. Microbiol.*, **21**, 43-54.

Bao, K. and Cohen, C.N. (2004). Reverse transcriptase activity innate to DNA polymerase I and DNA topoisomerase I proteins of *Streptomyces* telomere complex. *Proc. Natl. Acad. Sci. U S A.*, **In press.**

Bao, K. and Cohen, S.N. (2001). Terminal proteins essential for the replication of linear plasmids and chromosomes in Streptomyces. *Genes Dev.*, **15**, 1518-27.

Bao, K. and Cohen, S.N. (2003). Recruitment of terminal protein to the ends of *Streptomyces* linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication. *Genes Dev.*, **17**, 774-85.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., J., H., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., L., L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. and Hopwood, D.A. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, **417**, 141-147.

Bhattarcharjee, J.K. (1985). α-Aminoadipate pathway for the biosynthesis of lysine in lower eukaryotes. *Crit. Rev. Microbiol.*, **12**, 131-151.

Bibb, M.J., Findlay, P.R. and Johnson, M.W. (1984). The relationship between base composition and codon-usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene*, **30**. 157-166.

Bignell, D.R., Lau, L.H., Colvin, K.R. and Leskiw, B.K. (2003). The putative anti-anti-sigma factor BldG is post-translationally modified by phosphorylation in *Streptomyces coelicolor*. *FEMS Microbiol. Lett.*, **225**, 93-99.

Bignell, D.R., Warawa, J.L., Strap, J.L., Chater, K.F. and Leskiw, B.K. (2000). Study of the *bld*G locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. *Microbiology*, **146**, 2161-73.

Bignell, D.R.D., Tahlan, K., Colvin, K.R., Jensen, S.E. and Leskiw, B.K. (2004). Expression of *cca*R, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus*, is dependent on *bld*G. *Antimicrob. Agents Chemother.*, **In press.**

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **24**, 1513-1523.

Blackshear, P.J. (1984). Systems for polyacrylamide gel electrophoresis. *Methods Enzymol.*, **104**.

Box, S.J. (1977). Clavulanic acid and its salts by fermentation. *Chemical Abstracts*, **87**, 4031e.

Bradford, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of proteins utilizing the principle of protein dye-binding. *Anal. Biochem.*, **72**, 248-254.

Brana, A.F., Manzanal, M.B. and Hardisson, C. (1982). Mode of cell wall growth of *Streptomyces antibioticus*. *FEMS Microbiol. Lett.*, **13**, 231-235.

Brana, A.F., Paiva, N. and Demain, A.L. (1986). Pathways and regulation of ammonium assimilation in *Streptomyces clavuligerus*. *J Gen. Microbiol.*, **132**, 1305-1317.

Brannigan, J.A., Dodson, G., Duggleby, H., Moody, P.C.E., Smith, J.L., Tomchick, D.R. and Murzin, A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature*, **378**, 416-419.

Brewer, S.J., Taylor, P.M. and Turner, M.K. (1980). An adenosine triphosphate-dependent carbamoylphosphate-3-hydroxymethylcephem O-carbamoyltransferase from *Streptomyces clavuligerus*. *Biochem. J.*, **185**, 555-64.

Brown, A.G., Butterworth, G., Cole, M., Hanscombe, G., Hood, J.D., Reading, C. and Robinson, G.N. (1976). Naturally occurring β-lactamase inhibitors with antibacterial activity. *J. Antibiot.*, **29**, 668-669.

Brown, D., Evans, J.R. and Fletton, R.A. (1979). Structures of three novel β–lactams isolated from *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.*, **1979**, 282-283.

Brown, K.L., Wood, S. and Buttner, M.J. (1992). Isolation and characterization of the major vegetative RNA polymerase of *Streptomyces coelicolor* A3(2); renaturation of a sigma subunit using GroEL. *Mol. Microbiol.*, **6**, 1133-9.

Burton, K. (1957). A study of the conditions and mechanism of the diphenylamine reaction for the colorimetric estimation of deoxyribonucleic acid. *Biochem. J.*, **62**, 315-323.

Buttner, M.J., Chater, K.F. and Bibb, M.J. (1990). Cloning, disruption, and transcriptional analysis of three RNA polymerase sigma factor genes of *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, 172, 3367-78.

Buttner, M.J. and Lewis, C.G. (1992). Construction and characterization of *Streptomyces coelicolor* A3(2) mutants that are multiply deficient in the nonessential *hrd*-encoded RNA polymerase sigma factors. *J. Bacteriol.*, **174**, 5165-7.

Buttner, M.J., Smith, A.M. and Bibb, M.J. (1988). At least three different RNA polymerase holoenzymes direct transcription of the agarase gene (*dag*A) of *Streptomyces coelicolor* A3(2). *Cell*, **52**, 599-607.

Caines, M.E.C., Elkins, J.M., Hewitson, K.S. and Schofield, C.J. (2004). Crystal structure and mechanistic implications of N^2-(2-Carboxyethyl)arginine synthase, the first enzyme in the clavulanic acid biosynthesis pathway. *J. Biol. Chem.*, **279**, 5685-5692.

Calcutt, M.J. and Schmidt, F.J. (1992). Conserved gene arrangement in the origin region of the *Streptomyces coelicolor* chromosome. *J. Bacteriol.*, **174**, 3220-6.

Cashel, M., Gentry, D.R., Hernandez, V.J. and Vinella, D. (1996). The stringent response. In Neidhart, F.C. (ed.) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology Press, Washington, DC, pp. 1458-1496.

Chaconas, G. and van de Sande, J.H. (1980). 5'-^{32}P labeling of RNA and DNA restriction fragments. *Methods Enzymol.*, **65**, 75-85.

Chakraburtty, R. and Bibb, M. (1997). The ppGpp synthetase gene (*rel*A) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J. Bacteriol.*, **179**, 5854-61.

Challis, G.L. and Hopwood, D.A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomcyes species*. *Proc. Natl. Acad. Sci. U S A*, **100**, 14555-14561.

Champness, W.C. and Chater, K.F. (1994). Regulation and Integration of Antibiotic Production and Morphological Differentiation in *Streptomyces* spp. In Piggot, P.J., Moran Jr., C.P. and Youngman, P. (eds.), *Regulation of Bacterial Differentiation*. American Society for Microbiology, Washington, D.C., pp. 61-93.

Chater, K.F. (1993). Genetics of differentiation in *Streptomyces*. *Annu. Rev. Microbiol.*, **47**, 685-713.

Chater, K.F. and Bibb, M.J. (1997). Regulation of bacterial antibiotic production. In Kleinkauf, H. and von Dohren, H. (eds.), *Biotechnology*. VCH, Weinheim, Germany, Vol. 6, pp. 247-270.

Chater, K.F. and Losick, R. (1996). The mycelial life-style of *Streptomyces coelicolor* A3(2) and its relatives. In Shapiro, J.H. and Dworkin, M. (eds.), *Bacteria as multicellular organisms*. Oxford University Press, New York, pp. 149-182.

Chater, K.F. and Merrick, M.J. (1979). Streptomycetes. In Parish, J.H. (ed.) *Developmental Biology of Prokaryotes*. Blackwell Scientific Publications, Ltd., Oxford, pp. 93-114.

Chen, C.W. (1996). Complications and implications of linear bacterial chromosomes. *Trends Genet.*, **12**, 192-196.

Coque, J.J., Perez-Llarena, F.J., Enguita, F.J., Fuente, J.L., Martin, J.F. and Liras, P. (1995). Characterization of the *cmc*H genes of *Nocardia lactamdurans* and *Streptomyces clavuligerus* encoding a functional 3'-hydroxymethylcephem O-carbamoyltransferase for cephamycin biosynthesis. *Gene*, **162**, 21-27.

Cox, M.M. (1983). The FLP protein of the yeast 2-μm plasmid: Expression of a eukaryotic genetic recombination system in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A.*, **80**, 4223-4227.

Cunin, R., Glansdorff, N., Pierard, A. and Stalon, V. (1986). Biosynthesis and metabolism of arginine in bacteria. *Microbiol. Rev.*, **50**, 314-352.

Datsenko, K.A. and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U S A.*, **97**, 6640-6645.

de la Fuente, A., Martin, J.F., Rodriguez-Garcia, A. and Liras, P. (2004). Two proteins with ornithine acetyltransferase activity show different functions in *Streptomyces clavuligerus*: Oat2 modulates clavulanic acid biosynthesis in response to arginine. *J. Bacteriol.*, **186**, 6501-6507.

Demain, A.L. (2000). Small bugs, big business: The economic power of the microbe. *Biotechnol. Adv.*, **18**, 499-514.

Demain, A.L. and Elander, R.P. (1999). The beta-lactam antibiotics: past, present, and future. *Antonie van Leeuwenhoek*, **75**, 5-19.

Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Comm.*, **23**, 641-646.

Doran, J.L., Leskiw, B.K., Aippersbach, S. and Jensen, S.E. (1990a). Isolation and characterization of a β-lactamase-inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. *J. Bacteriol.*, **172**, 4909-18.

Doran, J.L., Leskiw, B.K., Petrich, A.K., Westlake, D.W. and Jensen, S.E. (1990b). Production of *Streptomyces clavuligerus* isopenicillin N synthase in *Escherichia coli* using two-cistron expression systems. *J. Ind. Microbiol.*, **5**, 197-206.

Egan, L.A., Busby, R.W., IwataReuyl, D. and Townsend, C.A. (1997). Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. *J. Am. Chem. Soc.*, **119**, 2348-2355.

Elander, R.P. (2003). Industrial production of β-lactam antibiotics. *Appl. Microbiol. Biotechnol.*, **61**, 385-392.

Elder, K. (1998). The characterization of an open reading frame involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Department of Biological Sciences. University of Alberta, Edmonton.

Elkins, J.M., Clifton, I.J., Hernandez, H., Doan, L.X., Robinson, C.V., Schofield, C.J., Hewitson, K.S. and Paradkar, A.S. (2002). Oligomeric structure of proclavaminic acid amidinohydrolase: evolution of a hydrolytic enzyme in clavulanic acid biosynthesis. *Biochem. J.*, **366**, 423-434.

Elkins, J.M., Kershaw, N.J. and Schofield, C.J. (2004). X-ray crystal structure of ornithine acetyltransferase from the clavulanic acid biosynthetic gene cluster. *Biochem. J.*, **In press.**

Elson, S.W. (1979). Studies on the biosynthesis of clavulanic acid. I. Incorporation of ^{13}C-labelled precursors. *J. Antibiot.*, **32**, 1125-1129.

Elson, S.W., Baggaley, K.H., Davison, M., Fulston, M., Nicholson, N.H., Risbridger, G.D. and Tyler, J.W. (1993a). The identification of three new biosynthetic intermediates and one further biosynthetic enzyme in the clavulanic acid pathway. *J. Chem. Soc. Chem. Commun.*, **1993**, 1212-1214.

Elson, S.W., Baggaley, K.H., Fulston, M., Nicholson, N.H., Tyler, J.W., Edwards, J., Holms, H., Hamilton, I. and Mousdale, D.M. (1993b). Two novel arginine derivatives from a mutant of *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.*, **1993**, 1211-1212.

Elson, S.W., Baggaley, K.H., Sime, J.T. and Woronieck, S.R. (1987). Isolation of two novel β-lactams and a novel dioxygenase cycling enzyme from *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.*, **1987**, 1736-1738.

Elson, S.W., Gillett, J., Nicholson, N.H. and Tyler, J.W. (1988). *N*-acyl derivatives of clavaminic acid produced by a mutant of *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.*, **1988**, 979-980.

Elson, S.W. and Oliver, R.S. (1978). Studies on the biosynthesis of clavulanic acid. I. Incorporation of ^{13}C-labelled precursors. *J. Antibiot.*, **31**, 586-92.

Eoyang, L. and Silverman, P.M. (1984). Purification and subunit composition of acetohydroxy acid synthase I from *Escherichia coli* K-12. *J. Bacteriol.*, **157**, 184-189.

Evans, M. and Dyson, P. (1993). Pulsed-field gel electrophoresis of *Streptomyces lividans* DNA. *Trends Genet.*, **9**, 72.

Fernandez-Moreno, M.A., Caballero, J.L., Hopwood, D.A. and Malpartida, F. (1991). The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bld*A tRNA gene of *Streptomyces*. *Cell*, **66**, 769-80.

Fernandez-Moreno, M.A., Martin-Triana, A.J., Martinez, E., Niemi, J., Kieser, H.M., Hopwood, D.A. and Malpartida, F. (1992). *aba*A, a new pleiotropic regulatory locus for antibiotic production in *Streptomyces coelicolor*. *J. Bacteriol.*, **174**, 2958-67.

Fleming, A. (1929). On the antibacterial action of *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *Brit. J. Exp. Path.*, **10**, 226-236.

Floriano, B. and Bibb, M. (1996). *afs*R is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **21**, 385-396.

Folcher, M., Gaillard, H., Nguyen, L.T., Nguyen, K.T., Lacroix, P., Bamas-Jacques, N., Rinkel, M. and Thompson, C.J. (2001). Pleiotropic functions of a *Streptomyces pristinaespiralis* autoregulator receptor in development, antibiotic biosynthesis, and expression of a superoxide dismutase. *J. Biol. Chem.*, **276**, 44297-44306.

Foulstone, M. and Reading, C. (1982). Assay of amoxicillin and clavulanic acid, the components of Augmentin, in biological fluids with high-performance liquid chromatography. *Antimicrob. Agents Chemother.*, **22**, 753-62.

Fuente, J.L., Rumbero, A., Martin, J.F. and Liras, P. (1997). δ-1-Piperideine-6-carboxylate dehydrogenase, a new enzyme that forms α-aminoadipate in *Streptomyces clavuligerus* and other cephamycin C producing actinomycetes. *Biochem. J.*, **327**, 59-64.

Fujii, T., Gramajo, H.C., Takano, E. and Bibb, M.J. (1996). *red*D and *act*II-ORF*4*, pathway-specific regulatory genes for antibiotic production in *Streptomyces coelicolor* A3(2), are transcribed in vitro by an RNA polymerase holoenzyme containing sigma HrdD. *J. Bacteriol.*, **178**, 3402-3405.

Galperin, M.Y. (2004). Bacterial signal transduction network in a genomic perspective. *Environ. Microbiol.*, **6**, 552-567.

Gevers, D., Vandepoele, K., Simillion, C. and Peer, Y.V. (2004). Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends in Microbiol.*, **12**, 148-154.

Gish, W. and States, D.J. (1993). Identification of protein coding regions by database similarity search. *Nat. Genet.*, **3**, 266-272.

Gowrishankar, J. and Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? *Mol. Microbiol.*, **54**, 598-603.

Gramajo, H.C., Takano, E. and Bibb, M.J. (1993). Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Mol. Microbiol.*, **7**, 837-845.

Gravius, B., Glocker, D., Pigac, J., Pandza, K., Hranueli, D. and Cullum, J. (1994). The 387 kb linear plasmid pPZG101 of *Streptomyces rimo*sus and its interactions with the chromosome. *Microbiology*, **140**, 2271-2277.

Gray, D.I., Gooday, G.W. and Prosser, J.I. (1990). Apical hyphal extension in *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.*, **136**, 1077-1084.

Guex, N. and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Gust, B., Challis, G.L., Fowler, K., Kieser, T. and Chater, K.F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U S A*, **100**, 1541-1546.

Gutman, A.L., Ribon, V. and Boltanski, A. (1985). Incorporation of β-hydroxyproprionate into the β-lactam residues of clavulanic acid. *J. Chem. Soc. Chem. Commun.*, **22**, 1627-1630.

Hardisson, C. and Manzanal, M.B. (1976). Ultrastructural studies of sporulation in *Streptomyces*. *J. Bacteriol.*, **127**, 1443-1454.

Hardisson, C., Manzanal, M.B., Salas, J.A. and Suarez, J.E. (1978). Fine structure, physiology and biochemistry of arthrospore germination in *Streptomyces antibioticus*. *J. Gen. Microbiol.*, **105**, 203-214.

He, J., Magarvey, N., Piraee, M. and Vining, L.C. (2001). The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology*, **147**, 2817-2829.

Higgens, C.E. and Kastnar, R.E. (1971). *Streptomyces clavuligerus* sp. nov. a β-lactam antibiotic producer. *Int. J. Syst. Bacteriol.*, **21**, 326-331.

Hill, C.M., Pang, S.S. and Duggleby, R.G. (1997). Purification of *Escherichia coli* acetohydroxyacid synthase isoenzyme II and reconstitution of active enzyme from its individual pure subunits. *Biochem. J.*, **327**, 891-898.

Hill, C.W. (1999). Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. *Res. Microbiol.*, **150**, 665-674.

Hodgson, J.E., Fosberry, A.P., Rawlinson, N.S., Ross, H.N., Neal, R.J., Arnell, J.C., Earl, A.J. and Lawlor, E.J. (1995). Clavulanic acid biosynthesis in *Streptomyces clavuligerus*: gene cloning and characterization. *Gene*, **166**, 49-55.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, K., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M. and Schrempf, H. (1985). *Genetic manipulation of Streptomyces: a Laboratory Manual*. John Innes Foundation, Norwich, United Kingdom.

Horinouchi, S. (2003). AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). *J. Ind. Microbiol. Biotechnol.*, **30**, 462-467.

Horinouchi, S. and Beppu, T. (1992a). Autoregulatory factors and communication in Actinomycetes. *Annu. Rev. Microbiol.*, **96**, 377-392.

Horinouchi, S. and Beppu, T. (1992b). Regulation of secondary metabolism and cell differentiation in Streptomyces: A-factor as a microbial hormone and the AfsR protein as a component of a two-component regulatory system. *Gene*, **115**, 167-172.

Horinouchi, S. and Beppu, T. (1994). A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. *Mol. Microbiol.*, **12**, 859-864.

Horinouchi, S., Kito, M., Nishiyama, M., Furuya, K., Hong, S.K., Miyake, K. and Beppu, T. (1990). Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). *Gene*, **95**, 49-56.

Howarth, T.T., Brown, A.G. and King, T.J. (1976). Clavulanic acid, a novel β-lactam from *Streptomyces clavuligerus*; X-ray crystal structure and analysis. *J. Chem. Soc. Chem. Commun.*, 266-267.

Hoyt, S. and Jones, G.H. (1998). *rel*A is required for actinomycin production in *Streptomyces antibioticus*. *J. Bacteriol.*, **181**, 3824-3829.

Hutchings, M.I., Hoskisson, P.A., Chandra, G. and Buttner, M.J. (2004). Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). *Microbiology*, **150**, 2795-2806.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.*, **21**, 526-531.

Ishikawa, J. and Hotta, K. (1999). FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. *FEMS Microbiol. Lett.*, **174**, 251-253.

Ishizuka, H., Horinouchi, S., Kieser, H.M., Hopwood, D.A. and Beppu, T. (1992). A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. *J. Bacteriol.*, **174**, 7585-7594.

Jensen, S.E., Alexander, D.C., Paradkar, A.S. and Aidoo, K.A. (1993). Extending the β-lactam biosynthetic gene cluster in *Streptomyces clavuligerus*. In Baltz, R.H., Hegeman, G.D. and Skatrud, P.L. (eds.), *Industrial Microorganisms: Basic and Applied Molecular Genetics*. American Society of Microbiology, Washington, D.C., pp. 169-176.

Jensen, S.E. and Demain, A.L. (1995). Beta-lactams. In Vining, L.C. and Sttuard, C. (eds.) *Genetics and Biochemistry of Antibiotic Production*. Butterworth-Heinemann, Boston, pp. 239-268.

Jensen, S.E., Elder, K.J., Aidoo, K.A. and Paradkar, A.S. (2000). Enzymes catalyzing the early steps of clavulanic acid biosynthesis are encoded by two sets of paralogous genes in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.*, **44**, 720-726.

Jensen, S.E., Leskiw, B.K., Vining, L.C., Aharonowitz, Y., Westlake, D.W. and Wolfe, S. (1986). Purification of isopenicillin N synthetase from *Streptomyces clavuligerus*. *Can. J. Microbiol.*, **32**, 953-8.

Jensen, S.E., Paradkar, A.S., Mosher, R.H., Anders, C., Beatty, P.H., Brumlik, M., Griffin, A. and Barton, B. (2004a). Five additional genes are involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.*, **48**, 192-202.

Jensen, S.E. and Paradkar, A.S. (1999). Biosynthesis and molecular genetics of clavulanic acid. *Antonie van Leeuwenhoek*, **75**, 125-133.

Jensen, S.E., Westlake, D.W., Bowers, R.J. and Wolfe, S. (1982a). Cephalosporin formation by cell-free extracts from *Streptomyces clavuligerus*. *J. Antibiot.*, **35**, 1351-1360.

Jensen, S.E., Westlake, D.W. and Wolfe, S. (1982b). Cyclization of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine to penicillins by cell-free extracts of *Streptomyces clavuligerus*. *J. Antibiot.*, **35**, 483-490.

Jensen, S.E., Wong, A., Griffin, A. and Barton, B. (2004b). *Streptomyces clavuligerus* has a second copy of the proclavaminate amidinohydrolase gene. *Antimicrob. Agents Chemother.*, **48**, 514-520.

Jin, W., Ryu, Y., Kang, S.G., Kim, S.K., Saito, N., Ochi, K., Lee, S.H. and Lee, K.J. (2004). Two *rel*A/*spo*T homologous genes are involved in the morphological and physiological differentiation of *Streptomyces clavuligerus*. *Microbiology*, **150**, 1485-1493.

Jones, G.H., Paget, M.S.B., Chamberlin, L. and Buttner, M.J. (1997). Sigma-E is required for the production of the antibiotic actinomycin in *Streptomyces antibioticus*. *Mol. Microbiol.*, **23**, 169-178.

Kawamura, Y., Kondo, E., Matsumoto, K., Tsuji, N. and Tawara, K. (1987a). Clavam antibiotic CA-31. *Chemical Abstracts*, **106**, 212565v.

Kawamura, Y., Kondo, E., Matsumoto, K., Tsuji, N. and Tawara, K. (1987b). Clavam antibiotics CA-146A and 146B. *Chemical abstracts*, **107**, 76085e.

Kershaw, N.J., McNaughton, H.J., Hewitson, K.S., Hernandez, H., Griffin, J., Hughes, C., Greaves, P., Barton, B., Robinson, C.V. and Schofield, C.J. (2002). ORF6 from the clavulanic acid gene cluster of *Streptomyces clavuligerus* has ornithine acetyltransferase activity. *Eur. J. Biochem.*, **269**, 2052-2059.

Khaleeli, N., Li, R.F. and Townsend, C.A. (1999). Origin of the β-lactam carbons in clavulanic acid from an unusual thiamine pyrophosphate-mediated reaction. *J. Am. Chem. Soc.*, **121**, 9223-9224.

Kholkov, A.S., Tovarova, I.I., Borisova, L.N., Pliner, S.A., Schevchenko, I.I., Kornitskaka, E.Y., Ivkina, N.S. and Rapoport, I.A. (1967). A-factor responsible for the production of streptomycin by a mutant strain of *Actinomyces streptomycini*. *Dokl. Akad. SSSR*, **177**, 232-235.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000). Practical *Streptomyces* genetics. John Innes Foundation, Norwich, England.

Kim, H.S., Lee, Y.J., Lee, C.K., Choi, S.U., Yeo, S.-H., Hwang, Y.I., Yu, T.S., Kinoshita, H. and Nihira, T. (2004). Cloning and characterization of a gene encoding the γ-butyrolactone autoregulator receptor from *Streptomyces clavuligerus*. *Arch. Microbiol.*, **182**, 44-50.

Kinashi, H. (1994). Linear plasmids from actinomycetes. *Actinomycetologica*, **8**, 87-96.

Kinashi, H., Mori, E., Hatani, A. and Nimi, O. (1994). Isolation and characterization of linear plasmids from lankacidin-producing *Streptomyces* species. *J. Antibiot.*, **47**, 1447-1455.

Kinoshita, H., Tsuji, T., Ipposhi, H., Nihira, T. and Yamada, Y. (1999). Characterization of binding sequences for butyrolactone autoregulator receptors in streptomycetes. *J. Bacteriol.*, **181**, 5075-5080.

Kirby, K.S., Fox-Carter, E. and Guest, M. (1967). Isolation of deoxyribonucleic acid and ribosomal ribonucleic acid from bacteria. *Biochem. J.*, **104**, 258-262.

Kirby, R., Wright, L.F. and Hopwood, D.A. (1975). Plasmid determined antibiotic synthesis and resistance in *Streptomyces coelicolor*. *Nature*, **254**, 265-267.

Kitano, K., Kintaka, K. and Katamoto, K. (1979). Clavulanic acid production by *Streptomyces katsurahamanus*. *Chemical Abstracts*, **90**, 119758b.

Kormanec, J. and Farkasovsky, M. (1993). Differential expression of principal sigma factor homologues of *Streptomyces aureofaciens* correlates with the developmental stage. *Nucleic Acids Res.*, **21**, 3647-3652.

Kormanec, J., Farkasovsky, M. and Potuckova, L. (1992). Four genes in *Streptomyces aureofaciens* containing a domain characteristic of principal sigma factors. *Gene*, **122**, 63-70.

Korn-Wendisch, F. and Kutzner, H.J. (1992). The family Streptomycetaceae. In Balows, A., Truper, H.G., Dworkin, M., Harder, W. and Schleifer, K.H. (eds.), *The Prokaryotes*. Springer-Verlag, New York, pp. 921-995.

Kovacevic, S. and Miller, J.R. (1991). Cloning and sequencing of the beta-lactam hydroxylase gene (*cef*F) from *Streptomyces clavuligerus*: gene duplication may have led to separate hydroxylase and expandase activities in the actinomycetes. *J. Bacteriol.*, **173**, 398-400.

Kovacevic, S., Tobin, M.B. and Miller, J.R. (1990). The β-lactam biosynthesis genes for isopenicillin N epimerase and deacetoxycephalosporin C synthetase are expressed from a single transcript in *Streptomyces clavuligerus*. *J. Bacteriol.*, **172**, 3952-3958.

Kovacevic, S., Weigel, B.J., Tobin, M.B., Ingolia, T.D. and Miller, J.R. (1989). Cloning, characterization, and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. *J. Bacteriol.*, **171**, 754-760.

Kyung, Y.S., Hu, W.S. and Sherman, D.H. (2001). Analysis of temporal and spatial expression of the CcaR regulatory element in the cephamycin C biosynthetic pathway using green fluorescent protein. *Mol. Microbiol.*, **40**, 530-541.

Larson, J.L. and Hershberger, C.L. (1986). The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. *Plasmid*, **15**, 199-209.

Lawlor, E.J., Baylis, H.A. and Chater, K.F. (1987). Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). *Genes Dev.*, **1**, 1305-1310.

Lawrence, G.C. and Lilly, G. (1980). United Kingdom Patent No. 1,571,888.

Leblond, P., Fischer, G., Francou, F., Berger, F., Guerineau, M. and Decaris, B. (1996). The unstable region of *Streptomyces ambofaciens* includes 210 kb terminal inverted repeats flanking the extremities of the linear chromosomal DNA. *Mol. Microbiol.*, **19**, 261-271.

Leblond, P., Redenbach, M. and Cullum, J. (1993). The physical map of the *Streptomyces lividans* 66 genome and comparison with that of the related strain *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, **175**, 3422-3429.

Lee, P.C., Umeyama, T. and Horinouchi, S. (2002). *afs*S is a target of AfsR, a transcriptional factor with ATPase activity that globally controls secondary metabolism in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **43**, 1413-1430.

Leskiw, B.K., Bibb, M.J. and Chater, K.F. (1991a). The use of a rare codon specifically during development. *Mol. Microbiol.*, **5**, 2861-2867.

Leskiw, B.K., Lawlor, E.J., Fernandez-Abalos, J.M. and Chater, K.F. (1991b). TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces mutants*. *Proc. Natl. Acad. Sci. U S A*, **88**, 2461-2465.

Leskiw, B.K., Mah, R., Lawlor, E.J. and Chater, K.F. (1993). Accumulation of *bld*A-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, **175**, 1995-2005.

Lezhava, A.L., Mizukami, T., Kajitani, T., Kameoka, D., Redenbach, M., Shinkawa, H., Nimi, O. and Kinashi, H. (1995). Physical map of the linear chromosome of *Streptomyces griseus*. *J. Bacteriol.*, **177**, 6492-6498.

Li, R., Khaleeli, N. and Townsend, C.A. (2000). Expansion of the clavulanic acid gene cluster: identification and in vivo functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. *J. Bacteriol.*, **182**, 4087-4095.

Lin, Y.S., Kieser, H.M., Hopwood, D.A. and Chen, C.W. (1993). The chromosomal DNA of *Streptomyces lividans* 66 is linear. *Mol. Microbiol.*, **10**, 923-933.

Liras, P. (1999). Biosynthesis and molecular genetics of cephamycins. Cephamycins produced by actinomycetes. *Antonie van Leeuwenhoek*, **75**, 109-124.

Liras, P., Rodriguez-Garcia, A. and Martin, J.F. (1998). Evolution of the clusters for β-lactam antibiotics: a model for evolutive combinatorial assembly of new β-lactams. *Int. Microbiol.*, **1**, 271-278.

Lorenzana, L.M., Perez-Redondo, R., Santamarta, I., Martin, J.F. and Liras, P. (2004). Two oligopeptide-permease-encoding genes in the clavulanic acid cluster of *Streptomyces clavuligerus* are essential for production of the β-Lactamase inhibitor. *J. Bacteriol.*, **186**, 3431-3438.

Macfarlane, G. (1984). *Alexander Flemming: The man and the myth.* Harvard University Press, Cambridge, MA.

Madduri, K., Stuttard, C. and Vining, L.C. (1989). Lysine catabolism in *Streptomyces* spp. is primarily through cadaverine: beta-lactam producers also make alpha-aminoadipate. *J. Bacteriol.*, **171**, 299-302.

Madduri, K., Stuttard, C. and Vining, L.C. (1991). Cloning and location of a gene governing lysine epsilon-aminotransferase, an enzyme initiating beta-lactam biosynthesis in *Streptomyces* spp. *J. Bacteriol.*, **173**, 985-988.

Malmberg, L.H., Hu, W.S. and Sherman, D.H. (1993). Precursor flux control through targeted chromosomal insertion of the lysine epsilon-aminotransferase (*lat*) gene in cephamycin C biosynthesis. *J. Bacteriol.*, **175**, 6916-69124.

Manzanal, M.B. and Hardisson, C. (1978). Early stages of arthrospore maturation in *Streptomyces*. *J. Bacteriol.*, **133**, 293-297.

Marck, C. (1988). 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res.*, **16**, 1829-1836.

Marcos, A.T., Gutierrez, S., Diez, B., Fernandez, F.J., Oguiza, J.A. and Martin, J.F. (1995). Three genes *hrd*B, *hrd*D and *hrd*T of *Streptomyces griseus* IMRU 3570, encoding sigma factor-like proteins, are differentially expressed under specific nutritional conditions. *Gene*, **153**, 41-48.

Marsh, E.N., Chang, M.D. and Townsend, C.A. (1992). Two isozymes of clavaminate synthase central to clavulanic acid formation: cloning and sequencing of both genes from *Streptomyces clavuligerus*. *Biochemistry*, **31**, 12648-12657.

Massova, I. and Mobashery, S. (1997). Molecular basis for interactions between β-lactam antibiotics and β-lactamases. *Acc. Chem. Res.*, **30**, 162-168.

Mathews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999). Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *J. Mol. Biol.*, **288**, 911-940.

Matsumoto, A., Hong, S.K., Ishizuka, H., S., H. and Beppu, T. (1994). Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species by a eukaryotic-type protein kinase. *Gene*, **146**, 47-56.

Matsumoto, A., Ishizuka, H., Beppu, T. and Horinouchi, S. (1995). Involvement of a small ORF downstream of the *afs*R gene in the regulation of secondary metabolism in *Streptomyces coelicolor* A3(2). *Actinomycetologica*, **9**, 37-43.

McCarthy, A.J. and Williams, S.T. (1992). Actinomycetes as agents of biodegradation in the environment-a review. *Gene*, **115**, 189-192.

McNaughton, H.J., Thirkettle, J.E., Zhang, Z.H., Schofield, C.J., Jensen, S.E., Barton, B. and Greaves, P. (1998). β-lactam synthetase: Implications for β-lactamase evolution. *Chem. Commun.*, **1998**, 2325-2326.

Mellado, E., Lorenzana, L.M., Rodriguez-Saiz, M., Diez, B., Liras, P. and Barredo, J.L. (2002). The clavulanic acid biosynthetic cluster of *Streptomyces* *clavuligerus*: genetic organization of the region upstream of the *car* gene. *Microbiology*, **148**, 1427-1438.

Mendez, C., Brana, A.F., Manzanal, M.B. and Hardisson, C. (1985). Role of substrate mycelium in colony development in *Streptomyces*. *Can. J. Microbiol.*, **31**, 446-450.

Merrick, M.J. (1976). A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. *J. Gen. Microbiol.*, **96**, 299-315.

Miguelez, E.M., Garcia, M., Hardisson, C. and Manzanal, M.B. (1994). Autoradiographic study of hyphal growth during aerial mycelium development in *Streptomyces antibioticus*. *J. Bacteriol.*, **176**, 2105-2107.

Miller, I.M., Stapley, E.O. and Chailet, L. (1962). Production of synnemantin B by the member of the genus *Streptomyces*. *Bacteriol. Proc.*, **49**, 32.

Miller, M.T., Bachmann, B.O., Townsend, C.A. and Rosenzweig, A.C. (2001). Structure of β-lactam synthetase reveals how to synthesize antibiotics instead of asparagine. *Nat. Struct. Biol.*, **8**, 684-689.

Miyadoh, S. (1993). Research on antibiotic screening in Japan over the last decade: a producing microorganisms approach. *Actinomycetologica*, **7**, 100-106.

Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, S., F., Yamada, K. and Kinashi, H. (2003). The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Mol. Microbiol.*, **48**, 1502-1510.

Mosher, R.H., Paradkar, A.S., Anders, C., Barton, B. and Jensen, S.E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.*, **43**, 1215-24.

Motamedi, H., Shafiee, A. and Cai, S.J. (1995). Integrative vectors for heterologous gene expression in *Streptomyces* spp. *Gene*, **160**, 25-31.

Musialowski, M.S., Flett, F., Scott, G.B., Hobbs, G., Smith, C.P. and Oliver, S.G. (1994). Functional evidence that the principal DNA replication origin of the *Streptomyces coelicolor* chromosome is close to the *dna*A-*gyr*B region. *J. Bacteriol.*, **176**, 5123-5125.

Nagarajan, R., Boeck, L.D., Gorman, M., Hamill, R.L., Higgens, C.E., Hoehn, M.M., Stark, W.M. and Whitney, J.G. (1971). Beta-lactam antibiotics from *Streptomyces*. *J. Am. Chem. Soc.*, **93**, 2308-2310.

Nageli, H.U., Loosli, H.R. and Nassbaumer, A. (1986). Clavamycins, new clavam antibiotics from two variants of *Streptomyces hygroscopicus*. II. Isolation and structures of clavamycins A, B and C from *Streptomyces hygroscopicus* NRRL 15899. *J. Antibiot.*, **39**, 516-524.

Narva, K.E. and Feitelson, J.S. (1990). Nucleotide sequence and transcriptional analysis of the *red*D locus of *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, **172**, 326-333.

Netolitzky, D.J., Wu, X., Jensen, S.E. and Roy, K.L. (1995). Giant linear plasmids of β-lactam antibiotic producing *Streptomyces*. *FEMS Microbiol. Lett.*, **131**, 27-34.

Nicholson, N.H., Baggaley, K.H., Cassels, R., Davison, M., Elson, S.W., Fulston, M., Tyler, J.W. and Woroniecki, S.R. (1994). Evidence that the immediate biosynthetic precursor of clavulanic acid is its N-aldehyde analog. *J. Chem. Soc. Chem. Commun.*, **1994**, 1281-1282.

Ochi, K. (1986). Occurrence of the stringent response in *Streptomyces* sp. and its significance for the initiation of morphological and physiological differentiation. *J. Gen. Microbiol.*, **132**, 2621–2631.

O'Connor, T.J., Kanellis, P. and Nodwell, J.R. (2002). The *ram*C gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol. Microbiol.*, **45**, 45-57.

OH, S. and Chater, K.F. (1997). Denaturation of circular or linear DNA facilitates targeted integrative transformation of *Streptomyces coelicolor* A3(2): possible relevance to other organisms. *J. Bacteriol.*, **179**, 122-127.

O'Keefe, D.P. and Harder, P.A. (1991). Occurrence and biological function of cytochrome P450 monooxygenases in actinomycetes. *Mol. Microbiol.*, **5**, 2099-2105.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y. and Hattori, M. (2001). Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U S A*, **98**, 12215-12220.

Onaka, H., Anido, N., Nihara, T., Yamada, Y., Beppu, T. and Horinouchi, S. (1995). Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*. *J. Bacteriol.*, **177**, 6083-6092.

Onishi, Y., Kameyama, S., Onaka, H. and Horinouchi, S. (1999). The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*:

identification of a target gene of the A-factor receptor. *Mol. Microbiol.*, **34**, 102-111.

Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J. and Buttner, M.J. (1999). Evidence that the extracytoplasmic function sigma factor sigma-E is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J Bacteriol*, **181**, 204-211.

Pandza, K., Pfalzer, G., Callum, J. and Haranueli, D. (1997). Physical mapping shows that the unstable oxytetracycline gene cluster of *Streptomyces rimosus* lies close to one end of the linear chromosome. *Microbiology*, **143**, 1493-1501.

Pandza, S., Biukovic, G., Paravic, A., Dadbin, A., Cullum, J. and Hranueli, D. (1998). Recombination between the linear plasmid pPZG101 and the linear chromosome of *Streptomyces rimosus* can lead to exchange of ends. *Mol. Microbiol.*, **28**, 1165-1176.

Paradkar, A., Trefzer, A., Chakraburtty, R. and Stassi, D. (2003). *Streptomyces* genetics: a genome perspective. *Crit. Rev. Biotechnol.*, **23**, 1-27.

Paradkar, A.S., Aidoo, K.A. and Jensen, S.E. (1998). A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Mol. Microbiol.*, **27**, 831-843.

Paradkar, A.S., Aidoo, K.A., Wong, A. and Jensen, S.E. (1996). Molecular analysis of a β-lactam resistance gene encoded within the cephamycin gene cluster of *Streptomyces clavuligerus*. *J. Bacteriol.*, **178**, 6266-6274.

Paradkar, A.S. and Jensen, S.E. (1995). Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *J. Bacteriol.*, **177**, 1307-1314.

Paradkar, A.S., Jensen, S.E. and Mosher, R.H. (1997). Comparative genetics and molecular biology of β-lactam biosynthesis. In Strohl, W.J. (ed.) *Biotechnology of Antibiotics*. Marcel Dekker, Inc., New York, pp. 241-277.

Peitsch, M.C. (1995). Protein modeling by E-mail. *Bio/Technology*, **13**, 658-660.

Perez-Llarena, F., Martin, J.F., Galleni, M., Coque, J.J., Fuente, J.L., Frere, J.M. and Liras, P. (1997a). The *bla* gene of the cephamycin cluster of *Streptomyces clavuligerus* encodes a class A β-lactamase of low enzymatic activity. *J. Bacteriol.*, **179**, 6035-6040.

Perez-Llarena, F.J., Liras, P., Rodriguez-Garcia, A. and Martin, J.F. (1997b). A regulatory gene *(cca*R*)* required for cephamycin and clavulanic acid production in

Streptomyces clavuligerus: amplification results in overproduction of both β-lactam compounds. *J. Bacteriol.*, **179**, 2053-2059.

Perez-Llarena, F.J., Rodriguez-Garcia, A., Enguita, F.J., Martin, J.F. and Liras, P. (1998). The *pcd* gene encoding piperideine-6-carboxylate dehydrogenase involved in biosynthesis of α-aminoadipic acid is located in the cephamycin cluster of *Streptomyces clavuligerus*. *J. Bacteriol.*, **180**, 4753-4756.

Perez-Redondo, R., Rodriguez-Garcia, A., Martin, J.F. and Liras, P. (1998). The *cla*R gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. *Gene*, **211**, 311-321.

Perez-Redondo, R., Rodriguez-Garcia, A., Martin, J.F. and Liras, P. (1999). Deletion of the pyc gene blocks clavulanic acid biosynthesis except in glycerol-containing medium: evidence for two different genes in formation of the C3 unit. *J. Bacteriol.*, **181**, 6922-6928.

Peter, H., Rabenrohst, J., Rohl, F. and Zahner, H. (1985). Recent advances in chemotherapy. In Ishigamt, J. (ed.). University of Tokyo Press, Tokyo, pp. 237-238.

Petrich, A.K., Leskiw, B.K., Paradkar, A.S. and Jensen, S.E. (1994). Transcriptional mapping of the genes encoding the early enzymes of the cephamycin biosynthetic pathway of *Streptomyces clavuligerus*. *Gene*, **142**, 41-48.

Petrich, A.K., Wu, X., Roy, K.L. and Jensen, S.E. (1992). Transcriptional analysis of the isopenicillin N synthase-encoding gene of *Streptomyces clavuligerus*. *Gene*, **111**, 77-84.

Petrickova, K. and Petricek, M. (2003). Eukaryotic-type protein kinases in *Streptomyces coelicolor*: variations to a common theme. *Microbiology*, **149**, 1609-1621.

Pruess, D.L. and Kellett, M. (1983). Ro 22-5417, a new clavam antibiotic from *Streptomyces clavuligerus*. I. Discovery and biological activity. *J. Antibiot.*, **36**, 208-212.

Redenbach, M., Ikeda, K., Yamaski, M. and Kinashi, H. (1998). Cloning and physical mapping of the *Eco*RI fragments from the giant linear plasmid SCP1. *J. Bacteriol.*, **180**, 2796-2799.

Roach, P.L., Clifton, I.J., Fulop, V., Harlos, K., Barton, G.J., Hajdu, J., Andersson, I., Schofield, C.J. and Baldwin, J.E. (1995). Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature*, **375**, 700-704.

Rodriguez-Garcia, A., de La Fuente, A., Perez-Redondo, R., Martin, J.F. and Liras, P. (2000). Characterization and expression of the arginine biosynthesis gene cluster of *Streptomyces clavuligerus*. *J. Mol. Microbiol. Biotechnol.*, **2**, 543-550.

Romero, J., Liras, P. and Martin, J.F. (1984). Dissociation of cephamycin and clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Appl. Microbiol. Biotechnol.*, **20**, 318-325.

Romero, J., Liras, P. and Martin, J.F. (1986). Utilization of ornithine and arginine as specific precursors of clavulanic acid. *Appl. Environ. Microbiol.*, **52**, 892-897.

Salowe, S.P., Krol, W.J., Iwata-Reuyl, D. and Townsend, C.A. (1991). Elucidation of the order of oxidations and identification of an intermediate in the multi-step clavaminate synthase reaction. *Biochemistry*, **30**, 2281-2292.

Salowe, S.P., Marsh, E.N. and Townsend, C.A. (1990). Purification and characterization of clavaminate synthase from *Streptomyces clavuligerus*: an unusual oxidative enzyme in natural product biosynthesis. *Biochemistry*, **29**, 6499-6508.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Samson, S.M., Dotzalf, J.E., Slisz, M.L., Becker, G.W., van Frank, R.M., Veal, L.E., Yeh, W., Miller, J.R., Queener, S.W. and Ingolia, T.D. (1987). Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. *Biotechnology*, **5**, 1207-1216.

Sanger, F., S. Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Sanraku-Ocean Co. Ltd. (1981). Clavulanic acid. *Chemical Abstracts*, **94**, 137803z.

Santamarta, I., Rodriguez-Garcia, A., Perez-Redondo, R., Martin, J.F. and Liras, P. (2002). CcaR is an autoregulatory protein that binds to the *cca*R and *cef*D-*cmc*I promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. *J. Bacteriol.*, **184**, 3106-3113.

Sawai, R., Suzuki, A., Takano, Y., Lee, P.C. and Horinouchi, S. (2004). Phosphorylation of AfsR by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). *Gene*, **9**, 53-61.

Scheu, A.K., Martinez, E., Soliveri, J. and Malpartida, F. (1997). *aba*B, a putative regulator for secondary metabolism in *Streptomyces*. *FEMS Microbiol. Lett.*, **147**, 29-36.

Schrempf, H., Bujard, H., Hopwood, D.A. and Gobel, W. (1975). Isolation of covalently closed circular deoxyribonucleic acid from *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, **121**, 416-421.

Schwecke, T., Aharonowitz, Y., Palissa, H., von Dohren, H., Kleinkauf, H. and van Liempt, H. (1992). Enzymatic characterization of the multifunctional enzyme δ-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase from *Streptomyces clavuligerus*. *Eur. J. Biochem.*, **205**, 687-694.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.*, **31**, 3381-3385.

Seki, C., Den, K., Marunaka, T., Miyake, Y., Minami, Y., Kajitani, A. and Ishada, N. (1987). Antitumor and antifungus G0069C substance produced by *Streptomyces* species. JP 61, 268, 285. *Chemical Abstracts*, **106**, 194774f.

Shinkawa, H., Hatada, Y., Okada, M., Kinashi, H. and Nimi, O. (1995). Nucleotide sequence of a principal sigma factor gene (*hrd*B) of *Streptomyces griseus*. *J. Biochem.*, **118**, 494-499.

Skatrud, P.L., Schwecke, T., Liempt, H.v. and Tobin, M.B. (1997). Advances in the molecular genetics of β-lactam antibiotic biosynthesis. In Kleinkauf, H. and Döhren , H.v. (eds.), *Biotechnology*. Wiley-VCH, Berlin, Vol. 7, pp. 247-270.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, **98**, 503-517.

Stackebrandt, E., Rainey, F.A. and Ward-Rainey, N.L. (1997). Proposal for a new hierarchic classification system, Actinobacteria classis nov. *Int. J. Syst. Bacteriol.*, **47**, 479-491.

- **Stapley, E.O., Jackson, M., Harnandex, S., Zimmerman, S.B., Currie, S.A., Mochalis, S., Mahta, J.M., Woodruff, H.B. and Hendlin, D. (1972).** Caphamycins a new family of β-lactam antibiotics. 1. Production by actinomyces, including *Streptomyces lactamdurans* sp. n. *Antimicrob. Agents Chemother.*, **2**, 122-131.

Strauch, E., Takano, E., Baylis, H.A. and Bibb, M.J. (1991). The stringent response in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **5**, 289-298.

Strohl, W.R. (1992). Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.*, **20**, 961-974.

Strynadka, N.C., Jensen, S.E., Johns, K., Blanchard, H., Page, M., Matagne, A., Frere, J.M. and James, M.N. (1994). Structural and kinetic characterization of a β-lactamase-inhibitor protein. *Nature*, **368**, 657-660.

Stuttard, C. (1982.). Temperate phages of *Streptomyces venezuelae*: lysogeny and host specificity shown by phages SV1 and SV2. *J. Gen. Microbiol.*, **128**, 115-121.

Stutzman-Engwall, K.J., Otten, S. and Hutchinson, C.R. (1992). Regulation of secondary metabolism in *Streptomyces* Spp. and overproduction of daunorubicin in *Streptomyces peucetius*. *J. Bacteriol.*, **174**, 144-154.

Sun, J., Kelemen, G.H., Fernandez-Abalos, J.M. and Bibb, M.J. (1999). Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3(2). *Microbiology*, **145**, 2221-2227.

Sunohara, T., Jojima, K., Tagami, H., Inada, T. and Aiba, H. (2004). Ribosome stalling during translation elongation induces cleavage of mRNA being translated in *Escherichia coli*. *J. Biol. Chem.*, **279**, 15368–15375.

Tahlan, K., Anders, C. and Jensen, S.E. (2004a). The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. *J. Bacteriol.*, **186**, 6286-6297.

Tahlan, K., Park, H.-U. and Jensen, S.E. (2004b). Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Can. J. Microbiol.*, 803-810.

Tahlan, K., Park, H.-U., Wong, A., Beatty, P.H. and Jensen, S.E. (2004c). Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.*, **48**, 930-939.

Takano, E., Gramajo, H.C., Strauch, E., Andres, N., White, J. and Bibb, M.J. (1992). Transcriptional regulation of the *red*D transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **6**, 2797-2804.

Tanaka, K., Shiina, T. and Takahashi, H. (1988). Multiple principal sigma factor homologs in eubacteria: identification of the "rpoD box". *Science*, **242**, 1040-1042.

Tang, L., Grimm, A., Zhang, Y.X. and Hutchinson, C.R. (1996). Purification and characterization of the DNA-binding protein DnrI, a transcriptional factor of daunorubicin biosynthesis in *Streptomyces peucetius*. *Mol. Microbiol.*, **22**, 801-813.

Thai, W., Paradkar, A.S. and Jensen, S.E. (2001). Construction and analysis of β-lactamase-inhibitory protein (BLIP) non-producer mutants of *Streptomyces clavuligerus*. *Microbiology*, **147**, 325-335.

Tipper, D.J. and Strominger, J.L. (1965). Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc. Natl. Acad. Sci. U S A*, **54**, 1133-1141.

Towbin, H.T., Stalhelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U S A*, **76**, 4350-4354.

Townsend, C.A. and Ho, M. (1985a). Biosynthesis of clavulanic acid: Origin of the C5 unit. *J. Am. Chem. Soc.*, **107**, 1065-1066.

Townsend, C.A. and Ho, M. (1985b). Biosynthesis of clavulanic acid: Origin of the C3 unit. *J. Am. Chem. Soc.*, **107**, 1066-1068.

Townsend, C.J. and Krol, W.J. (1988). The role of molecular oxygen in clavulanic acid biosynthesis: evidence for a bacterial oxidative deamination. *J. Chem. Soc. Chem. Commun.*, **1988**, 1234-1236.

Trepanier, N.K., Jensen, S.E., Alexander, D.C. and Leskiw, B.K. (2002). The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bld*A mutant. *Microbiology*, **148**, 643-656.

Umbarger, H. (1978). Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.*, **47**, 533-606.

Umeyama, T. and Horinouchi, S. (2001). Autophosphorylation of bacterial serine/threonine kinase, AfsK, is inhibited by KbpA, an AfsK binding protein. *J. Bacteriol.*, **183**, 5506-5512.

Umeyama, T., Lee, P.C. and Horinouchi, S. (2002). Protein serine/threonine kinases in signal transduction for secondary metabolism and morphogenesis in *Streptomyces*. *Appl. Microbiol. Biotechnol.*, **59**, 419-425.

Umeyama, T., Lee, P.-C., Ueda, K. and Horinouchi, S. (1999). An AfsK/AfsR system involved in the response of aerial mycelium formation to glucose in *Streptomyces griseus*. *Microbiology*, **145**, 2281-2292.

Usui, S. and Yu, C.A. (1989). Purification and properties of isopenicillin N epimerase from *Streptomyces clavuligerus*. *Biochim. Biophys. Acta*, **999**, 78-85.

Valentine, B.P., Bailey, C.R., Doherty, A., Morris, J., Elson, S.W., Baggaley, K.H. and Nicholson, N.H. (1993). Evidence that arginine is a later metabolic intermediate than ornithine in the biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.*, **1993**, 1210-1211.

Vieira, J. and Messing, J. (1987). Production of single stranded plasmid DNA. *Methods Enzymol.*, **153**, 3-11.

Vujaklija, D., Horinouchi, S. and Beppu, T. (1993). Detection of an A-factor responsive protein that binds to the upstream activation sequence of *str*R, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. *J. Bacteriol.*, **175**, 2652-2661.

Vyazmensky, M., Sella, C., Barak, Z. and Chipman, D.E. (1996). Isolation and characterization of subunits of acetohydroxy acid synthase isozyme III and reconstitution of the holoenzyme. *Biochemistry*, **35**, 10339-10346.

Waksman, S.A. and Henrici, A.T. (1943). The nomenclature and classification of the actinomycetes. *J. Bacteriol.*, **46**, 337-341.

Walsh, C. (2003). *Antibiotics*. American Society for Microbiology Press, Washington, DC.

Wang, L., Tahlan, K., Kaziuk, T.L., Alexander, D.C. and Jensen, S.E. (2004). Transcriptional and translational analysis of the *cca*R gene from *Streptomyces clavuligerus*. *Microbiology*, 4137-4145.

Wanning, M., Zahnner, H., Krone, B. and Zeek, A. (1981). Ein neus anti-fungisches β-lactam antibioticum der clavam-reihe. *Tetrahedron Lett.*, **22**, 2539-2540.

Ward, J.M. and Hodgson, J.E. (1993). The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster' in three *Streptomyces*. *FEMS Microbiol. Lett.*, **110**, 239-242.

Ward, J.M., Janssen, G.R., Kieser, T., Bibb, M.J. and Buttner, M.J. (1986). Construction and characterization of a series of multi-copy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase gene from Tn5 as indicator. *Mol. Gen. Genet.*, **203**, 468-478.

Waxman, D.J. and Strominger, J.L. (1982). β-Lactam antibiotics: Biochemical modes of action. In Morin, R.B. and Gorman, M. (eds.), *Chemistry and biology of β-lactam antibiotics*. Academic Press, New York, Vol. 3, pp. 209-285.

Wehmeier, U.F. (1995). New multifunctional *Escherichia coli-Streptomyces* shuttle vectors allowing blue-white screening on X-Gal plates. *Gene*, **165**, 149-150.

Westpheling, J., Ranes, M. and Losick, R. (1985). RNA polymerase heterogeneity in *Streptomyces coelicolor*. *Nature*, **313**, 22-27.

White, J. and Bibb, M. (1997). *bld*A dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *J. Bacteriol.*, **179**, 627-633.

Wietzorrek, A. and Bibb, M. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol. Microbiol.*, **25**, 1181-1184.

Wildermuth, H. (1970). Development and organization of the aerial mycelium in *Streptomyces coelicolor*. *J. Gen. Microbiol.*, **60**, 43-50.

Williams, J.G. and Mason, P.J. (1985). *Hybridization in the analysis of RNA*. IRL Press, Oxford.

Wise, E.M.J. and Park, J.T. (1965). Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis. *Proc. Natl. Acad. Sci. U S A*, **54**, 75-81.

Wu, T.K., Busby, R.W., Houston, T.A., McIlwaine, D.B., Egan, L.A., Townsend, C.A. and Lode, H. (1995). Identification, cloning, sequencing, and overexpression of the gene encoding proclavaminate amidino hydrolase and characterization of protein function in clavulanic acid biosynthesis. *J. Bacteriol.*, **177**, 3714-3720.

Wu, W. (2003). Characterization of pSCL2, a giant linear plasmid in *Streptomyces clavuligerus*. Department of Biological Sciences. University of Alberta, Edmonton.

Wu, X. and Roy, K.L. (1993). Complete nucleotide sequence of a linear plasmid from *Streptomyces clavuligerus* and characterization of its RNA transcripts. *J. Bacteriol.*, **175**, 37-52.

Xiao, X., Bowers, R.J., Shin, H.-S., Wolfe, S. and Demain, A.L. (1991a). Non-radioactive assay and HPLC analysis of Cephalosporin C 7-α-methoxylation by cell-free extracts of *Streptomyces clavuligerus*. *Appl. Microbiol. Biotechnol.*, **35**, 793-797.

Xiao, X.F., Wolfe, S. and Demain, A.L. (1991b). Purification and characterization of cephalosporin 7 alpha-hydroxylase from *Streptomyces clavuligerus*. *Biochem. J.*, **280**, 471-474.

Yamasaki, M., Ikuto, Y., Ohira, A., Chater, K. and Kinashi, H. (2003). Limited regions of homology between linear and circular plasmids encoding methylenomycin biosynthesis in two independently isolated streptomycetes. *Microbiology*, **149**, 1351-1356.

Yamazaki, H., Tomono, A., Ohnishi, Y. and Horinouchi, S. (2004). DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. *Mol. Microbiol.*, **53**, 555-572.

Yang, C.C., Huang, C.H., Li, C.Y., Tsay, Y.G., Lee, S.C. and Chen, C.W. (2002). The terminal proteins of linear *Streptomyces* chromosomes and plasmids: a novel class of replication priming proteins. *Mol. Microbiol.*, **43**, 297-305.

Yang, M.C. and Losick, R. (2000). Cytological evidence for association of the ends of the linear chromosome in *Streptomyces coelicolor*. *J. Bacteriol.*, **183**, 5180-5186.

Zakrzewska-Czerwinska, J. and Schrempf, H. (1992). Characterization of an autonomously replicating region from the *Streptomyces lividans* chromosome. *J. Bacteriol.*, **174**, 2688-2693.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406-3415.